

高动态范围成像研究进展及展望

刘晓宁, 张乐, 朱策*

(电子科技大学信息与通信工程学院, 四川成都 611731)

摘要:高动态范围成像技术旨在还原真实场景的亮度与色彩信息,克服传统传感器成像中普遍存在的高光过曝与暗部细节丢失问题,已拓展至自动驾驶、虚拟现实/增强现实等领域。然而,动态场景下的伪影去除仍是其核心挑战。围绕上述关键问题,系统回顾了相关数据集与评价指标,全面梳理了重要研究进展,并深入剖析了成像模型缺陷及当前技术瓶颈的内在成因。从模型泛化能力、计算复杂度及推理时间等维度,比较分析了现有先进方法的性能差异。进一步结合前沿发展趋势,给出核心基础挑战、关键性能优化与前沿技术探索三个层面的关键研究方向,以期为相关学术研究与工程实践提供参考。

关键词:高动态范围成像;曝光融合;运动估计;Transformer;扩散模型

中图分类号:TP391.4 **文献标志码:**A **文章编号:**1001-2486(2026)03-052-22

Advances and prospects in high dynamic range imaging

LIU Xiaoning, ZHANG Le, ZHU Ce*

(School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: High dynamic range imaging aims to recover the luminance and color information of real-world scenes, thereby overcoming the common problems of highlight saturation and shadow detail loss in conventional sensor imaging. It has been extended to applications such as autonomous driving and virtual reality/augmented reality. However, artifact removal in dynamic scenes remains a central challenge. To address this issue, the related datasets and evaluation metrics were systematically reviewed, the major research advances were comprehensively summarized, and the inherent causes of imaging model deficiencies and the current technical bottlenecks were further analyzed. It also compared and analyzed the performance differences among existing state-of-the-art methods from the perspectives of model generalization ability, computational complexity, and inference time. Building on recent development trends, three levels of important research directions were further identified, namely fundamental core challenges, key performance optimization, and frontier technology exploration, with the aim of providing a useful reference for both academic research and engineering practice.

Keywords: high dynamic range imaging; exposure fusion; motion estimation; Transformer; diffusion model

现实世界中的光照场景拥有极其宽广的亮度范围,其动态范围往往跨越数个数量级^[1](如自然界中最低光照可以达到 10^{-5} cd/m²,最大可达 10^9 cd/m²),远超传统显示设备或成像传感器所能记录和再现的能力^[2]。这种记录能力与真实场景之间的差距,催生了对高动态范围(high dynamic range, HDR)成像技术的迫切需求。HDR成像旨在完整捕获、处理、存储并最终再现真实场景中丰富的亮度与色彩信息,以解决传统传感器

动态范围不足所导致的高光过曝与暗部细节丢失等固有问题^[3-5]。如图1所示,HDR成像技术能够清晰还原高光与阴影中的屋顶细节,然而低或标准动态范围(low/standard dynamic range, LDR/SDR)限制了色彩和亮度的范围。自1997年Debevec等^[6]在SIGGRAPH会议上奠定HDR辐射度校准技术基础以来,该领域历经20余年的发展,已从最初的实验室理论,演变为贯穿图像采集、处理、压缩、显示全链路的成熟技术体系^[7-9]。

收稿日期:2025-12-19

基金项目:国家自然科学基金-重点国际(地区)合作研究基金资助项目(62020106011);四川省自然科学基金重大基金资助项目(2025ZNSFSC0002)

第一作者:刘晓宁(1992—),男,甘肃白银人,博士研究生,E-mail:liuxiaoning2016@sina.com

*通信作者:朱策(1969—),男,四川自贡人,教授,博士,博士生导师,E-mail:eczhu@uestc.edu.cn

引用格式:刘晓宁,张乐,朱策.高动态范围成像研究进展及展望[J].国防科技大学学报,2026,48(3):52-73.

Citation: LIU X N, ZHANG L, ZHU C. Advances and prospects in high dynamic range imaging [J]. Journal of National University of Defense Technology, 2026, 48(3): 52-73.



图1 高动态范围与标准动态范围的比较

Fig.1 Comparison between HDR and SDR

近年来,在计算摄影与计算机视觉技术的共同推动下,HDR 成像的研究范畴持续深化与拓展:①研究重心已从早期的多曝光图像融合^[10-11](multi-exposure image fusion, MEF),延伸至 HDR 传感器设计^[12-14]、色调映射算子^[15-17](tone mapping operators, TMOs)、逆色调映射算子^[18-20](inverse/reverse tone mapping operators, i/rTMOs)、HDR 视频压缩^[21-23]与传输^[24-25]以及质量评价^[26-28]等全链路关键技术;②深度学习以及生成式人工智能、大模型等新兴技术的融入,为解决传统挑战(如动态场景运动伪影或“鬼影”消除)开辟了全新路径^[29-30]。这一发展趋势也推动 HDR 成像技术实现了从“物理扩展”到“算法创造”的核心跨越:从早期依赖多帧曝光堆栈^[11]的传统方法,逐步发展为结合运动补偿^[31]与硬件加速^[32]的高效方案,再到人工智能(artificial intelligence, AI)驱动的语义分割优化^[33]、端到端重建^[34],乃至当前基于扩散模型^[35]的细节生成与单帧 HDR 传感技术,动态范围覆盖能力、实时处理效率与场景适配性持续突破^[36-41]。

随着技术体系的成熟,HDR 成像已从专业摄影、影视工业与医疗影像等传统优势领域,成功拓展至自动驾驶、虚拟/增强现实及超高清视频产业等前沿应用,展现出日益广泛的产业应用价值与科学研究前景^[42]。而伴随 HDR10^[43]、杜比视界(Dolby vision)、HDR Vivid 等标准的逐步完善,以及 HDR10 + Advanced 等新型标准的演进,技术产业化生态不断成熟,但动态场景鬼影(ghosting)去

除、饱和区域内容恢复、噪声抑制、低功耗实时处理、跨设备一致性适配等核心挑战仍未完全解决^[44-49]。

鉴于此,本文以 HDR 成像的核心挑战为切入点,系统回顾该领域的代表性数据集与评价指标体系,全面梳理关键研究进展,深度剖析技术瓶颈的本质成因,并结合深度学习、计算摄影等前沿技术趋势,勾勒 HDR 成像领域的未来发展路径,为相关研究的推进与实际工程应用提供有益参考与借鉴。

1 高动态范围成像概述

1.1 成像管线

如图2所示,图像获取^[6,50]是将现实场景转化为数字图像的过程,核心是通过设备(如相机)捕捉场景辐射度、经信号转换与处理生成可存储/传输的数字图像。数字图像获取流程主要包括:

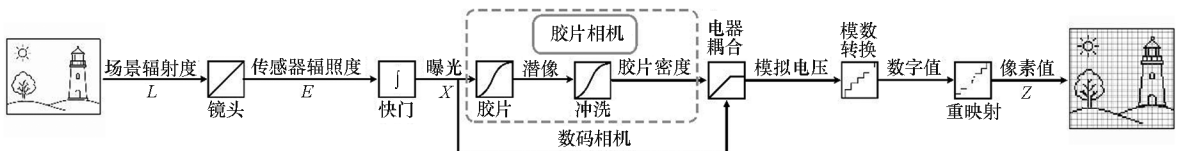
场景辐射捕获:场景的辐射度 L (如光照、物体反射)通过镜头聚焦,进入如胶片或镜头感光元件电荷耦合器件(charge-coupled device, CCD)或互补金属氧化物半导体(complementary metal-oxide semiconductor, CMOS)的表面形成传感器辐照度 E ,这是图像获取的“初始光信号输入”。

曝光控制:快门调节传感器的曝光时间,控制传感器辐照度的累积,生成传感器曝光 X ,曝光时间直接影响图像亮度与细节。

成像与信号转换:涵盖胶片相机与数字相机。其中,胶片相机是曝光后胶片经显影/冲洗生成潜像,再通过胶片密度转化为模拟信号;数字相机则是传感器(如 CCD)直接将光信号转换为模拟电压。

模数转换(analog-to-digital conversion, ADC):模拟信号(胶片密度/模拟电压)通过模数转换器转化为数字值——这是“模拟→数字”的关键步骤,决定图像的量化精度。

重映射:数字值经过重映射处理(如伽马校正、色彩空间转换等),最终生成像素强度值 Z ,即完成“可读化”与“标准化”。

图2 图像采集示意图^[6]Fig.2 Image acquisition schematic^[6]

1.2 问题定义及求解

1.2.1 动态范围及曝光

在介绍 HDR 成像物理模型之前,首先介绍 HDR 成像相关的度量,主要包括动态范围(dynamic range, DR)、曝光值(exposure value, EV)和感光度(ISO sensitivity)^[1],为统一表示成像相关参数,分别以 P_{DR} 、 P_{EV} 和 P_{ISO} 来指代。具体定义如下所示。

动态范围是指场景中亮度最大值与最小值的比值,亮度常用 cd/m^2 表示。实际场景中动态范围的线性比值通常较大,如 $10^{-6} \sim 10^8 \text{ cd}/\text{m}^2$,动态范围的比值可达到 10^{14} 。这种表示方式既不直观也不方便比较。为了将超大的线性比值“压缩”到易处理的数值范围内,常用以 10 为底的对数(单位为 dB)来表示:

$$P_{DR} = 20 \cdot \lg \frac{L_{\max}}{L_{\min}} \quad (1)$$

式中, L_{\max} 和 L_{\min} 分别为亮度最大值和最小值。这种对数转换的方式便于将较大的动态范围用简洁的“dB”数值来描述和比较。

在图像传感器领域,曝光值 P_{EV} 量化了拍摄场景的光线强度,由光圈系数 N 、快门速度(曝光时间 Δt)和感光度 P_{ISO} 共同决定:

$$P_{EV} = \log_2 \frac{N^2}{\Delta t} + \log_2 \frac{P_{ISO}}{100} \quad (2)$$

式中, P_{EV} 衡量的是“传感器最终接收到的等效曝光量”,而 P_{ISO} 决定了传感器对光的灵敏度。相同光圈和快门下, P_{ISO} 越高,传感器越灵敏,等效曝光量越多, P_{EV} 越大。

1.2.2 成像模型^[6]

假定在静态场景且同一视点下(即固定相机镜头不变)采集 P 幅不同曝光时长的图像 $\{I_k\}_{k=1}^P$,第 j 张照片的曝光时间为 Δt_j ,则相应像素点 i 的强度值记为 Z_{ij} 。由于场景是静态的,则每个像素点 i ($i=1, 2, \dots, M$, 其中 M 指像素点数量)对应的传感器辐照度 E_i 恒定。

1.2.3 模型建立

恢复 HDR 场景辐射度的关键是相机响应函数^[51](camera response function, CRF) f 的估计,根据互易假设,该函数将传感器辐照度 E_i 映射为像素强度值 Z_{ij} 且满足如下非线性映射关系:

$$Z_{ij} = f(E_i \Delta t_j) \quad (3)$$

式中,假设映射 f 单调且可逆,对其取自然对数并定义函数 $g = \ln f^{-1}$,则式(3)可转化为线性等式,即:

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (4)$$

从式(4)可以看出,所求未知量含有函数 g 与每个像素的辐照度对数 $\ln E_i$ 。无法直接从式(2)求未知量,主要原因为:①变量耦合且超/欠定;②实际拍摄场景存在噪声与误差;③缺乏物理约束。

1.2.4 模型求解

针对式(4)的局限性,引入“数据拟合项”即应对噪声,实现稳健求解;引入“平滑项”即物理约束,保证解的合理性。于是,得到如下最小化问题求解:

$$\min_{g, E_i} \sum_{i=1}^M \sum_{j=1}^P [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]^2 + \lambda \cdot \sum_{z=Z_{\min}+1}^{Z_{\max}-1} [g(z+1) - 2g(z) + g(z-1)]^2 \quad (5)$$

式中, λ 是权重参数,函数 g 的定义域为像素强度值的取值范围 $\{Z_{\min}, Z_{\min} + 1, \dots, Z_{\max}\}$,共有 $Z_{\max} - Z_{\min} + 1$ 个离散值。

通过奇异值分解来求解上述超定线性方程组,解得 g 后结合加权融合策略,即给位于响应函数中间区域的像素赋予更高权重,有效降低噪声和饱和影响,进而可还原传感器辐照度 E_i ,即:

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij}) [g(Z_{ij}) - \ln \Delta t_j]}{\sum_{j=1}^P w(Z_{ij})} \quad (6)$$

式中,权重函数 $w(\cdot)$ 为帽函数,可以弱化饱和和区域的影响。最后根据生成的 HDR 辐射图像像素强度值与场景真实辐射值成比例关系^[52],进而得到场景辐射度 L ^[6]。对于彩色图像,三通道独立求解响应函数,可通过校正光源,调整通道缩放因子保证色彩平衡,更多细节请参考文献[6]。

如图 3 所示,利用 Debevec 等^[6]提出的方法首先对不同曝光的 LDR 图像(该曝光序列总共含有 16 幅不同曝光的图像,为了便于可视化,仅列出 3 幅)进行响应函数估计,重建 HDR 场景辐射度,然后选择合适的色调映射方法^[53]对 HDR 图像可视化。

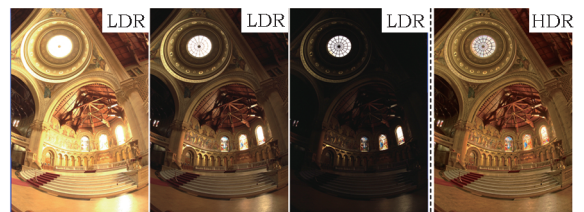


图 3 LDR 与 HDR 重建对比

Fig. 3 Comparison of LDR and HDR reconstruction

1.2.5 模型缺陷

Debevec 等^[6]建立了一个从多曝光图像到

HDR 辐射图重建的技术链路,通过恢复相机相应函数,将有限动态范围的图像融合为能准确反映场景真实亮度分布的 HDR 图像。该方法不仅解决了摄影领域长期存在的动态范围受限问题,更成为计算机图形学、计算机视觉和图像处理领域的重要基础技术,至今仍被广泛应用和研究。在 HDR 领域,该方法目前仍被用来构建动态场景的数据集标签^[54],但其存在以下主要问题:

- 1) 核心假设受限:依赖传感器互易性假设(极端曝光时间失效)和静态场景、光照稳定假设,动态场景或长时拍摄会导致偏差。
- 2) 数据采集要求严苛:需精准配准多曝光图像,抽样像素(方程求解时)需覆盖全灰度范围且低方差(主观性强)。
- 3) 权重参数 λ 经验化: λ 过小会导致 g 拟合

图像噪声(尤其暗部/亮部噪声); λ 过大则会过度光滑 g ,丢失场景的细节(如纹理、边缘)。

4) 噪声稳健性不足^[55-56]:当图像噪声(如传感器读出噪声、胶片颗粒噪声)较强时,会显著降低辐射度估计的精度。

5) 色彩与量化误差:三通道独立求解易导致色彩失衡,8 位/通道扫描设备会导致量化误差,影响饱和区域与暗部区域辐射度恢复精度。

6) 抽样与计算局限:仅用 50 个左右的抽样像素(易缺乏极端区间代表性),奇异值分解求解规模受限,无法利用全像素信息。

1.3 HDR 数据集

当前 HDR 数据的不同点主要体现在以下几个方面:曝光级、数据集规模、图像分辨率、图像质量和场景类型等。如表 1 所示,所整理的数据集包括:HDR 重建、视频 HDR 重建和 MEF。

表 1 现存 HDR 数据集比较
Tab.1 Comparison of existing HDR datasets

数据集	数据集规模	图像质量	标签	图像分辨率	场景类型
Sen 等 ^[57]	8	真实	×	1 350 像素 × 900 像素	多帧 HDR
Tursun 等 ^[42]	16	真实	×	1 024 像素 × 682 像素	多帧 HDR
Kalantari 等 ^[54]	89	真实	✓	1 500 像素 × 1 000 像素	多帧 HDR
Prabhakar 等 ^[58]	582	真实	✓	100 万 ~ 400 万像素	多帧 HDR
Hu 等 ^[59]	250	合成	✓/×*	多分辨率 ^a	多帧 HDR
Liu 等 ^[46]	281	真实	✓	多分辨率 ^b	多帧 HDR
Tel 等 ^[60]	144	合成	✓	1 496 像素 × 1 000 像素	多帧 HDR
Endo 等 ^[18]	1 043	合成	✓	512 像素 × 512 像素	单帧 HDR
Eilertsen 等 ^[61]	96	合成	✓	1 024 像素 × 768 像素	单帧 HDR
Liu 等 ^[62]	—	合成/真实	✓	1 536 像素 × 1 024 像素	单帧 HDR
Zhang 等 ^[63]	41 222	合成/真实	✓	128 像素 × 64 像素	单帧 HDR 全景
Nemoto 等 ^[64]	46	合成	✓	—	单帧 HDR 人眼跟踪
Kim 等 ^[65]	59 818	合成	✓	4K 超高分辨率	单帧 HDR + 超分辨
Chen 等 ^[66]	97	真实	✓	4 096 像素 × 2 168 像素	HDR 视频(静态)
Chen 等 ^[66]	284	真实	✓/×*	4 096 像素 × 2 168 像素	HDR 视频(动态)
Kalantari 等 ^[67]	9	真实	×	1 280 像素 × 720 像素	HDR 视频(动态)
Shu 等 ^[9]	500	真实	✓	6 000 像素 × 4 000 像素	HDR 视频(动态)
Ma 等 ^[68]	24	真实	×	1K ~ 2K	MEF(静态)
Cai 等 ^[69]	589	真实	✓	3K ~ 6K	MEF(静/动态)

注:“—”表示没有说明;“*”表示部分没有标签;a 表示分辨率包括 1 590 像素 × 1 166 像素和 1 500 像素 × 1 000 像素;b 表示分辨率包括 4 000 像素 × 3 000 像素、4 608 像素 × 3 456 像素、4 096 像素 × 3 072 像素。

单帧 HDR: 对于单帧 HDR 图像数据集, 最常见的方法是使用不同的 CRF 函数模拟 LDR 成像管线, 这些函数来自一组 CRF 数据库^[18, 62] 或者使用虚拟相机基于随机选择的相机曲线捕捉场景的多个随机区域^[61, 66, 70]。为了创建一个真实的单帧 HDR 数据集, 文献[62]首先通过固定三脚架来拍摄多幅不同曝光的 LDR 图像序列。然后使用 Photomatix(图像编辑软件)将 LDR 图像进行曝光堆栈, 以获取相应的 HDR 图像。

多帧 HDR: 当前多帧 HDR 的数据集主要包括: Sen 等^[57]、Tursun 等^[42]、Kalantari 等^[54]、Prabhakar 等^[58]、Hu 等^[59]、Liu 等^[46]和 Tel 等^[60]。其中, 以 Kalantari 等^[54]为代表的数据集构建过程最具代表性。对于动态场景数据集, 文献[54]根据 Debevec 等^[6]提出的方法进行 HDR 场景辐射度恢复。具体地, 采用两阶段曝光堆栈策略:

1) 用三脚架固定相机, 拍摄 3 幅不同曝光的静态场景 LDR 图像 $\{I_k\}_{k=1}^3$ (曝光序列间精确对齐), 相应的曝光时间为 $\{\Delta t_j\}_{j=1}^3$, 利用 Debevec 等^[6]提出的辐射度校准方法进行 HDR 场景重建, 进而生成标签。

2) 从曝光序列 $\{I_k\}_{k=1}^3$ 中移除 I_1 和 I_3 ; 只留中间帧 I_2 作为参考帧。

3) 根据 Δt_1 和 Δt_3 的时间来重新拍摄 2 幅新的 LDR 图像 I'_1 和 I'_3 。固定相机拍摄过程中, 主体可以运动或者也可以有其他动态物体, 然后将 I_2 与新的 I'_1 和 I'_3 构成新的动态场景的曝光序列。

上述两阶段曝光堆栈生成动态场景序列及标签的策略被后续相关 HDR 数据集构建的方法广泛采用, 如 Hu 等^[59]、Liu 等^[46]和 Tel 等^[60]。由于当前这些 HDR 数据集场景辐射度的恢复方法仍然依赖于 Debevec 等^[6]提出的方法, 所以 1.2.5 节讨论的模型缺陷在标签生成的过程中仍然存在。

此外, Liu 等^[45]指出当前构建的一些数据集存在明显的缺陷, 例如: ①不精确地输入, 所提供的曝光值与实际的曝光值不符^[59]; ②稠密的不明显的动态背景^[54], 如树枝、树叶、流云和流水等。这类稠密的动态背景在构建数据集的过程中不可避免, 由于无法精确对齐, 当前先进的 HDR 重建方法即使可以去除大范围的运动伪影, 也会遭受明显的背景伪影干扰; ③显著的失真, 场景内容存在扭曲^[58]; ④有限的场景范围, 如数据集基准^[54, 60]缺乏夜景, 尽管 Liu 等^[46]尝试构建夜景场景, 但 ISO 增益以及长曝光的设置, 夜景场景内容

中的噪声以及模糊比较显著。

1.4 HDR 相关任务

为了捕捉、处理和重现自然场景中极端黑暗或明亮区域的细节, HDR 成像技术受到了学术界和工业界的广泛关注, 正在逐渐占据 LDR/SDR 成像技术的应用市场。传统的 HDR 成像主要依赖于特殊的 HDR 相机来采集, 然而昂贵的设备成本限制了其广泛的应用。此外, 由于缺乏合适的 HDR 显示设备, HDR 相机记录的内容主要是通过 TMOs^[15, 71-74]来显示在普通的 LDR 显示器上。然而, 在摄影、电影制作、印刷和可视化领域, 将自然和人造场景复现到动态范围有限的显示设备上是一个具有挑战性的问题^[75]。因为 TMOs 通过降低图像中的色调值来实现 HDR 内容再现, 这可能导致图像的细节丢失, 甚至会呈现不可避免的内容结构失真^[76-77]。

除专业的 HDR 成像设备外, 最流行的 HDR 成像技术是利用特别设计的 HDR 成像算法将 LDR 的图像进行重建。当前, 根据输入图像的帧数将 HDR 成像方法分为两类^[42, 74], 即单帧 HDR 和多帧 HDR。前者是将单帧曝光的 LDR 图像重建为相应的 HDR 图像。这个过程也被称为逆(反向)色调映射算子^[18, 78-80](i/rTMOs), 它涉及重新创建丢失的信息, 并将可用信息扩展和适应到更高的位深度, 如从 16 bit 映射到 32 bit。iTMOs 主要集中在扩展动态范围, 以适应 LDR 内容到新的位深度。与多帧 HDR 相比, 虽然避免了运动伪影去除, 但是这类方法常用幻觉的方式来恢复饱和和区域的结构和纹理。在一定程度上, 这必然会导致空间位置上的结构与纹理不一致, 无法保证高保真重建。

多帧 HDR 算法旨在将相同场景下不同曝光的 LDR 图像合成一幅细节信息更丰富的 HDR 图像。与单帧曝光相比, 多帧曝光可以提供更丰富的细节信息, 所以当前消费级影像设备主要采用多帧曝光的融合方案设计, 尤其是智能手机中的 HDR 摄影技术的应用^[46, 81-82]。而与 MEF 相比, 多帧 HDR 确保了跨设备显示或渲染时具备更高的可扩展性与适应性, 同时保持卓越的视觉质量。这是因为 MEF 只是对曝光图像进行融合进而提高对比度, 但并不改变图像的位深度。

考虑到多帧 HDR 应用的广泛性, 本文主要回顾多帧 HDR 的研究进展、潜力与挑战。为了叙述简洁, 将多帧 HDR 记作 HDR。事实上, 在 HDR 领域, 通常 HDR 重建是指多帧情形。

1.5 HDR 图像质量评价

HDR 图像质量评价是衡量场景辐射度恢复精确性的关键环节,包括主观与客观两类方法。常用的客观评价可分为三类:像素级误差指标、结构相似性指标、视觉感知指标。

1.5.1 像素级误差指标

峰值信噪比 (peak signal-to-noise ratio, PSNR)^[83]:基于像素级均方误差 (mean squared error, MSE) 的全参考图像质量评价指标,单位为 dB。

设原始图像为 I ,重建或生成图像为 K ,二者尺寸均为 $H \times W$ (高度 \times 宽度),则 PSNR 定义为:

$$Q_{\text{PSNR}} = 10 \cdot \lg \left(\frac{I_{\text{MAX}}^2}{Q_{\text{MSE}}(I, K)} \right) \quad (7)$$

式中, I_{MAX} 为图像像素值的动态范围 (例如 8 位图像为 255), $Q_{\text{MSE}}(I, K)$ 为重建前与重建后图像之间的均方误差:

$$Q_{\text{MSE}}(I, K) = \frac{1}{HW} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} [I(i, j) - K(i, j)]^2 \quad (8)$$

PSNR 数值越高表示图像质量越好 (失真越小)。在 HDR 重建任务中,主要包括:PSNR- l 和 PSNR- μ 。其中,PSNR- l 用来评价在 HDR 线性域 (辐射域) 中,重建的 HDR 图像与真值 (ground truth, GT) 的差异;PSNR- μ 指将重建后的 HDR 图像以及 GT 首先进行 μ 律映射^[54],然后计算二者的差异。

μ 律映射是一种典型的色调映射函数,通过对数变换的非线性映射,将 HDR 图像的线性域亮度值 x 压缩到窄动态范围的“色调映射域”,即:

$$T(x) = \frac{\ln(1 + \mu x)}{\ln(1 + \mu)} \quad (9)$$

式中, $T(\cdot)$ 为色调映射算子, $\mu = 5\,000$ 。此外, PSNR-M 利用 MATLAB 中的内置 tonemap 函数作为色调映射来计算信噪比。

1.5.2 结构相似性指标

结构相似度 (structural similarity index measure, SSIM):用于衡量 2 幅图像的统计相似性^[84],包括亮度、对比度和结构信息。其中,亮度信息用于描述图像的平均亮度值,对比度信息用于刻画图像的变化程度,结构信息用于表征图像的纹理特征。

SSIM 的计算可以简化为:

$$Q_{\text{SSIM}}(I, K) = \frac{(2\mu_I\mu_K + C_1)(2\sigma_{IK} + C_2)}{(\mu_I^2 + \mu_K^2 + C_1)(\sigma_I^2 + \sigma_K^2 + C_2)} \quad (10)$$

式中, I 和 K 分别表示待比较的 2 幅图像, μ_I 和 μ_K 表示相应图像的均值, σ_I 和 σ_K 表示相应图像的标准差, σ_{IK} 表示图像 I 和图像 K 的协方差, C_1 和 C_2 是两个取值非常小的常数,避免分母为零和分子过大。

在 HDR 重建中,类似于 PSNR 在双域的计算,SSIM 也包括:SSIM- l 和 SSIM- μ 。

1.5.3 视觉感知指标

高动态范围可见性差异预测 (high dynamic range visible difference predictor, HDR-VDP)^[85]:基于校准的人眼视觉系统模型,预测 HDR 重建图像与参考图像之间的差异是否会被观察者感知。其中, HDR-VDP-2^[27] 进一步被设计为同时预测可见性差异和图像质量,并基于新的对比敏感度函数模型覆盖不同亮度条件。其核心优势在于模拟人眼早期视觉系统 (从眼内光散射到神经信号处理),可适配从夜间到白天的全亮度范围 (0.02 ~ 150 cd/m²),解决了传统指标 (如 PSNR、HDR-VDP^[85]) 仅适用于低亮度范围的局限。

该指标通过量化“平均观察者检测图像差异的概率” (可见性) 和“主观平均意见得分” (质量),为 HDR 成像、色调映射、图像压缩等算法提供客观评估依据,性能在低亮度场景显著优于 HDR-VDP,质量预测效果媲美甚至超越 MS-SSIM^[86]。其主要局限为仅关注亮度差异,并未考虑颜色和时间维度的影响。此外, HDR-VQM^[87] 用来评价 HDR 视频重建的质量。更多详情请参见文献^[85] 和^[87]。

基于深度学习的质量评价:上述评价指标都是依靠手动设计的特征来评价重建后的 HDR 图像。随着深度学习的发展,基于数据驱动的评价指标如雨后春笋般涌现^[88-92]。其中,文献^[91-92] 是非参考性的 HDR 评价指标,用来量化从神经网络中提取的底层和高级特征。

1.6 HDR 成像比赛

IEEE 国际计算机视觉与模式识别会议 (conference on computer vision and pattern recognition, CVPR) 的 workshop,即图像恢复和增强的新趋势 (new trends in image restoration and enhancement, NTIRE) 举行了三届与 HDR 相关的挑战赛,即 NTIRE-HDR-2021/2022^[93-94] 和 NTIRE 2025 Efficient Burst HDR and Restoration^[95]。例如,前两届 HDR 挑战赛分为两个赛道:单帧 HDR 和多帧 HDR。该比赛所采用的数据集来自先前的工作,具体地, Froehlich 等^[96] 使用一台专业的双相机装置构建了一个广泛的、多样化的、具有挑

战性的 HDR 视频数据集。输入的 LDR 图像通过一个像素测量模型^[55]来获取,其中包括多种噪声类型。关于 HDR 成像和噪声模型的更多细节以及数据集收集过程,请参见文献[93]。

2 高动态范围成像方法

图 4(数据来源于 Web of Science)呈现了 1995—2025 年间以 HDR 为主题的期刊文章数量

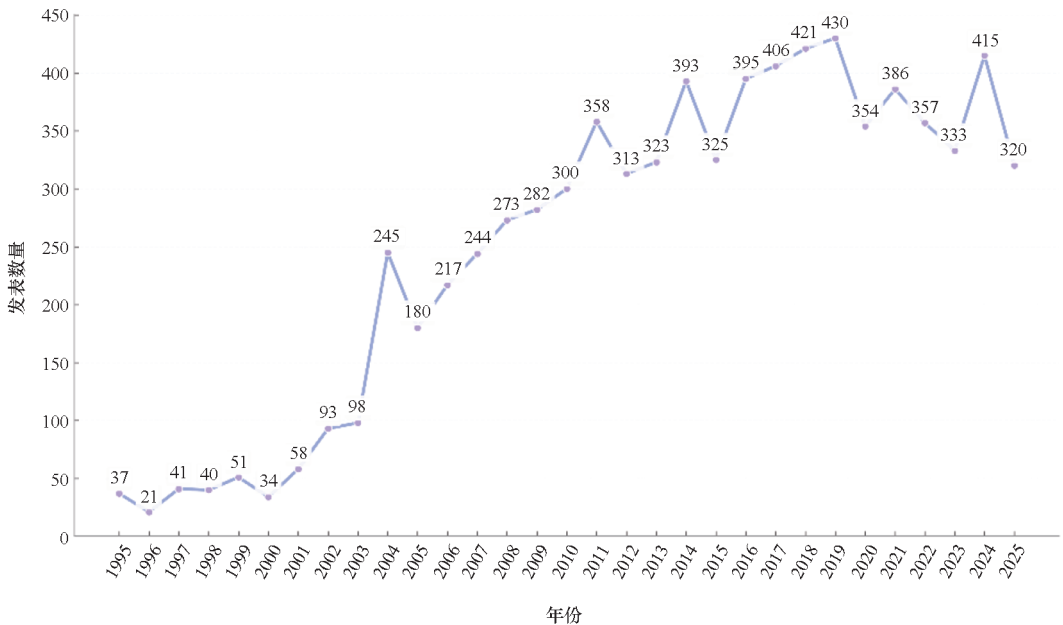


图 4 高动态范围成像为主题的期刊文章数量

Fig. 4 Number of journal articles on high dynamic range imaging

HDR 成像是从多幅不同曝光的 LDR 图像中,高效合成出一幅既无鬼影又能高保真还原场景细节的 HDR 图像。该领域的相关研究主要划分为两类技术路径:传统方法(手动设计特征)与深度学习驱动的方法。其核心逻辑是通过融合多幅曝光参数不同的 LDR 图像(欠曝光帧保留高光信息、过曝光帧记录暗部细节),实现对场景宽动态范围的完整覆盖。但在实际拍摄中,相机抖动(全局空间位移)或运动物体(局部位置变化)的存在,会导致不同曝光帧间的场景信息错位,直接重建后易产生明显的鬼/伪影。除抑制鬼影外,HDR 图像的保真度同样是核心要求;需避免高亮区域出现过饱和失真,同时完整保留其细节信息,以还原场景的真实视觉效果。

因此,如何在镜头抖动、物体大幅运动、严重遮挡及曝光变化(多曝光帧间区域存在饱和)的情况下,高效且准确地实现多帧融合是 HDR 重建的核心挑战。

为系统梳理 HDR 去鬼影技术的研究脉络,本文将现有方法划分为传统方法与基于深度学习的

的变化趋势:①该领域起步期(1995—2000年),文章数量整体偏低且波动平缓,多数年份在 30~40 篇之间,年均约 37 篇;②初步增长期(2001—2010年),发文量逐步抬升并出现波动;③快速发展期(2011—2025年),相关研究数量进入高频波动的高增长阶段,多次突破 350 篇,2018、2019 年分别达到 421、430 篇的峰值,直观体现了高动态范围成像领域的研究关注度与学术活力持续提升。

方法两大类。下文将围绕该分类框架,全面回顾近 30 年来 HDR 重建领域的关键技术进展。

2.1 传统方法

为了避免重建后 HDR 的运动伪影,传统意义上的 HDR 通过检测或消除鬼影的方式来解决该问题。根据处理方式的不同,传统方法主要分为三类^[36, 42, 54]:基于运动区域移除的方法、基于图像对齐的方法以及基于块优化的方法。

2.1.1 基于运动区域移除的方法

基于运动区域移除的方法认为不同曝光图像的大部分(背景)区域内容是静止的,只有小部分(前景)区域是动态的。显然,这类方法的关键问题是:如何根据输入 LDR 图像的像素来判断其所在的区域属性,即判断像素属于静止区域还是运动区域。大致流程为:①对输入图像进行全局对齐。②判断像素属性,如果像素属于静止区域,可以直接用传统的 HDR 方法进行融合;如果像素属于动态区域,则可以检测并移除该像素。

这些方法的区别主要体现在如何检测运动区域上。具体地,Khan 等^[97]计算给定像素是背景

的概率,并相应地分配权重。Jacobs 等^[98]通过计算曝光堆栈中不同图像的局部熵来检测运动区域的像素。Pece 和 Kautz^[99]为每个曝光图像计算中值阈值位图生成运动图。Zhang 和 Cham^[100]发现梯度的幅度可以表示像素的曝光质量,梯度的方向变化揭示运动物体。

另外一些方法是通过已知的相机响应函数来预测从一张曝光图像到另外一张曝光图像的颜色,并将其与原始像素颜色进行比较以检测运动区域^[101-103]。Heo 等^[104]通过计算非参考帧的每个像素到参考帧对应像素颜色的高斯加权距离为每个像素分配权重。Granados 等^[105]检测图像堆栈中一致的像素子集,然后解决标记问题以生成视觉上令人满意的 HDR 结果。此外,Zheng 等^[106]和 Li 等^[107]用双向方法检测不一致像素。秩最小化模型^[108-109]也被用于移除异常值并重建最终的 HDR 图像。

当输入曝光序列数量较少、存在动态背景(如水流)以及非刚性目标遮挡时,基于运动区域移除的方法无法处理 HDR 图像中的移动内容,因为这些方法只是简单地移除相应的像素。

2.1.2 基于图像对齐的方法

该类方法首先对齐多曝光图像,然后再将它们合并成一张 HDR 图像。若场景是平面刚体,即多曝光序列间的差异仅来自镜头抖动(比如轻微平移、旋转),没有物体运动或深度变化导致的局部错位,此时无须复杂的密集匹配或动态跟踪,用简单的几何变换^[110]、单应性矩阵估计^[111]、相关性匹配^[112]、相位跨相关性^[113]就能实现曝光序列的精准对齐。若非平面刚体场景(如相机抖动的同时有物体运动),面对这类场景,简单配准算法会失效,需要使用稠密但非刚性的对应关系,如单应性结合 SIFT^[104]、RANSAC^[114-115]、单应性结合光流^[116]等方法处理局部错位,避免配准后出现孔洞或鬼影。例如,Bogoni 等^[117]使用光流来估计局部运动以对齐输入图像。Kang 等^[116]使用光流变体^[118]来估计光流并提出了一种特定的 HDR 合并过程来移除对齐过程中的伪影。Jinno 和 Okuda^[119]将运动模糊建模为马尔可夫随机场以估计位移场。Zimmer 等^[120]通过由梯度和平滑项构成的最小化能量函数来计算光流。Hu 等^[121]通过使用 Hacohen 等^[122]提出的方法找到稠密的对应关系来对齐图像。Gallo 等^[114]提出了一种适用于较小幅度运动的估计方法。这些方法使用简单的合并方法来组合对齐的 LDR 图像。因此,无法有效避免大范围运动情况下的伪影。

2.1.3 基于块优化的方法

这类方法通过找到从参考帧到非参考帧的像素或图像块之间的对应关系,将非参考帧中有价值的像素/图像块信息转移到参考帧相应的区域;借助泊松方程^[102, 123]、图割^[124]、马尔可夫随机场^[105, 119]、贝叶斯^[125]等优化框架来执行对齐与重建。例如,Sen 等^[57]提出了基于块匹配^[126]的方法重建 HDR 图像,该方法处理的 LDR 图像是通过线性 CRF 相机获得的。其主要思想是通过循环迭代的方式来得到中间图像,该图像是从原始图像中搜索相似图像块并投票后填补获得,而不单是从估计的 HDR 图像中获得。Hu 等^[121]提出了类似的基于块匹配的优化框架,但将相机校准作为优化的一部分。

基于块的思路也被延伸到 MEF 任务中,尽管 HDR 与 MEF 的目标存在明显差异(前者输出场景辐射图,后者输出对比度平衡的低动态范围图像)。例如,Zheng 等^[106, 127]首先对两幅图像块利用强度映射函数来消除曝光所带来的亮度差异;然后利用图像块匹配的方法来恢复运动区域;最后去除曝光序列中的鬼影。Ma 等^[128-129]提出基于结构块分解的多曝光图像融合(structural patch decomposition based multi-exposure image fusion, SPD-MEF)方法,该方法对鬼影效果具有较强的稳健性。具体地,其将一个图像块分解为三个独立的部分,即信号强度、信号结构和平均强度;然后利用不同的度量标准得到三个理想的组件;最后将获得的三个组件融合后来重建 HDR 图像。对于运动的物体而言,SPD-MEF^[128]方法利用信号结构的方向信息来移除不一致的图像块。为了在融合时减少运动伪影,SPD-MEF 方法又额外引入强度映射函数^[130](intensity mapping function, IMF)来检查结构块间的一致性。考虑到 SPD-MEF^[128]方法计算耗时且存在光晕,Li 等^[131-132]分别提出多尺度分解的 SPD-MEF (multi-scale SPD-MEF, MSPD-MEF)和边缘保留的 SPD-MEF (edge-preserving SPD-MEF, ESPD-MEF)。

在传统意义上,虽然这些方法是 HDR 去鬼影领域的最新技术,但当参考图像具有大范围过饱和区域、遮挡区域、大幅度的运动物体时,基于图像块优化合成的方法往往会产生令人无法满意的结果。

2.2 基于深度学习的方法

得益于卷积神经网络(convolutional neural networks, CNNs)的层级非线性特征提取能力,近

年来基于 CNNs 的 HDR 重建方法^[36] (覆盖多曝光图像融合、单图 HDR 重建等典型任务) 已获得学术界与工业界的广泛关注。这类方法通过构建端到端深度网络, 学习 LDR 输入图像到 HDR 输出图像的特征映射关系, 借助局部感受野与逐层特征抽象能力, 实现基础的亮度信息融合与伪影抑制^[33, 133 - 134]。

随着深度学习架构的迭代, Transformer、扩散模型、大模型等技术也逐步拓展至 HDR 重建领域, 针对性地解决 CNNs 的局限:

1) Transformer: 凭借自注意力机制突破 CNNs 的局部感受野约束, 可精准捕捉图像全局尺度下的亮度依赖关系 (如大场景中高光天空与阴影建筑的亮度协同), 在复杂动态场景重建中, 有效提升跨区域亮度一致性, 减少大尺度场景下的运动伪影问题^[45, 81, 135]。

2) 扩散模型: 基于迭代去噪的生成式概率框架, 通过逐步逼近真实 HDR 图像的概率分布, 可从单幅 LDR 图像中恢复饱和区域的精细细节 (如过曝光源的纹理、暗部的轮廓), 同时抑制重建过程中的噪声与边缘伪影, 从而显著增强 HDR 图像的真实感与纹理丰富度^[35, 136 - 141]。

上述架构的持续演进, 正推动 HDR 重建技术向“高精度细节恢复、多场景泛化性、低功耗实时处理”的核心目标迈进。

2.2.1 基于光流对齐的方法

对于多帧 LDR 图像的对齐, 常用的方法是使用经典的光流算法^[142 - 144] 或当前的光流网络。将对齐后的 LDR 图像输入深度神经网络中, 以构建 HDR 图像。Kalantari 和 Ramamoorthi^[54] 将 CNNs 引入 HDR 重建领域。具体地, 其用经典的光流算法^[143] 将欠曝光和过曝光图像对齐到中等曝光的图像, 即参考帧; 对齐后的 LDR 图像序列作为输入, 由 HDR 图像的 GT 监督, 通过 CNNs 来学习 LDR 到 HDR 的映射, 从而重建相应的 HDR 图像。值得注意的是, 文献^[54] 还构建了一个目前使用很广泛的多帧 HDR 数据集, 其包括动态场景下不同曝光的 LDR 图像以及相应的无鬼影的 HDR 图像。Yan 等^[145] 改进了上述方法, 提出了基于多尺度 CNNs 的重建方法, 并将三个子网络的结果融合后获得相应的 HDR 图像。图 5 展示了基于光流对齐的处理流程。

然而, 传统的光流算法^[142 - 144] 时常会导致较大的对齐误差。于是, Peng 等^[146] 和 Prabhakar 等^[58] 分别使用先进的光流网络, 即 FlowNet^[147] 和 FlowNet 2.0^[148] 将非参考帧与参考帧进行对齐。

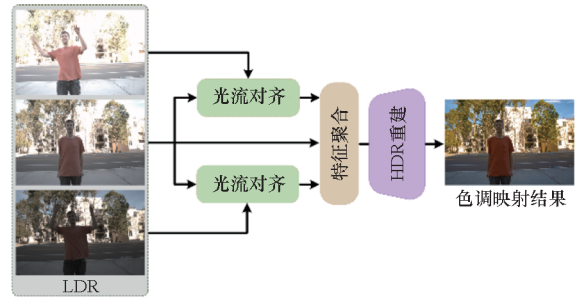


图 5 基于光流对齐的处理流程

Fig. 5 Pipeline based on optical flow alignment

Catley-Chandar 等^[149] 对曝光与光流对齐的不确定性进行建模, 提出 FlexHDR。相比先前的基于光流的方法, 文献^[149] 取得了显著的性能提升。虽然现有基于光流的方法有效, 但计算成本较高、延迟大、部署难。为此, Kong 等^[150] 提出 SAFNet, 其先提取金字塔特征, 根据共享编码器联合优化有效区域掩码与跨曝光运动, 借助金字塔光流逐级进行 HDR 聚合。在聚合的过程中, SAFNet 利用先进的光流网络生成掩码对初始融合系数进行重加权。此外, 针对长距离样本训练难的问题, Kong 等额外引入窗口分区裁剪策略。SAFNet 在当前基准数据集上取得了先进的性能, 且运行速度也优于先前基于 Transformer 的去鬼影方法^[60, 81]。

2.2.2 基于特征级联的方法

不同于上述光流对齐的方法, 特征级联的方法直接级联提取特征。当 LDR 曝光序列中存在大尺度的运动前景、遮挡以及过饱和的区域时, 上述的光流算法^[142 - 144] 或光流网络的对齐结果往往不够精确, 随着网络的传播, 其不可避免地积累了估计误差。为此, Wu 等^[29] 提出了 1 个自编码器框架, 包括 3 个编码器网络、1 个合成网络和 1 个解码器网络。编码器将不同曝光的 LDR 图像编码到潜在特征空间; 接下来, 编码后的 LDR 图像特征被输入合成网络中学习对齐特征; 最后, 解码器重建 HDR 图像。文献^[151] 中采用了类似的策略, 但与文献^[29] 不同的是, 其添加了 LDR 子网络来恢复对应的 HDR 图像。LDR 图像的循环重构增加了更多的约束, 有利于恢复 HDR 图像中的运动区域。

Chaudhari 等^[152] 提出了一个端到端的高动态范围图像信号处理管线, 其直接将不同曝光的原始彩色滤波阵列 (color filter array, CFA) 数据映射到 HDR 图像中。Metwaly 等^[153] 分别用 3 个编码器来提取 3 幅不同曝光下的 LDR 图像的特征。类似于文献^[29], 文献^[153] 将提取的特征级联

输入合成网络中。该方法的一个重要贡献是注意力掩码,旨在使网络关注具有显著运动的场景内容,以避免解码器产生鬼影效应。考虑到使用单独的编码器无法充分利用特征中的多尺度上下文信息,Niu 等^[154]提出多尺度 LDR 编码器来提取不同尺度下的视觉特征,类似于文献[145]的策略,所提取的特征通过残差学习在不同尺度上融合,然后被输入融合模块中。图6展示了基于特征级联的处理流程。

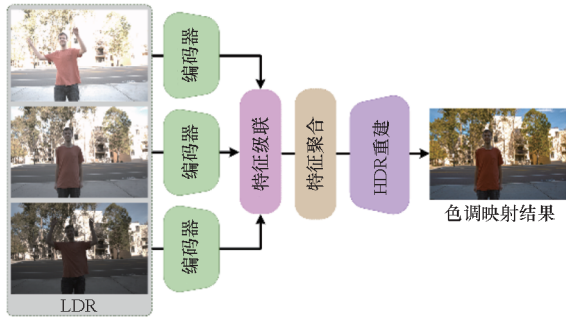


图6 基于特征级联的处理流程

Fig. 6 Pipeline based on feature concatenation

Tel 等^[60]提出 SCTNet, 直接对浅层特征进行级联,然后在合成阶段利用帧间与帧内注意力来实现对齐。尽管该方法相较于先前的方法表现出色,但通过实验可以观察到其仍然面临以下问题:①受架构冗余及自注意力二次计算复杂度的影响,模型存在显存制约,易引发“显存溢出”(out-of-memory, OOM)问题;②由于数据集多样性有限及模型领域适应性不足,其对未见过的域外场景泛化能力较弱;③推理速度较慢,无法适应移动端设备的需求。

2.2.3 基于相关性引导的方法

在实际摄影的过程中,场景中的物体运动形式可能是多种多样的且难以预计的。这使得基于光流的 HDR 重建方法很难抑制鬼影效应。最近的深度学习方法开始探讨如何确定不同曝光度和运动的低动态范围输入之间的相关性。使用注意力机制^[155]是解决这些问题的代表性方法。例如,Yan 等^[156]遵循文献[29]的流程,提出 AHDRNet,将注意力模块应用于编码阶段,以实现隐式的特征对齐,如图7所示。具体来说,注意力机制的特征提取网络,可以提取 LDR 图像中对参考帧具有互补性信息的特征,并抑制饱和、欠曝光和目标运动等低质量区域的特征;基于扩张残差稠密单元的融合网络,可以将 LDR 图像的特征进行有效的融合,扩大了网络的感受野从而关注远距离的运动信息。由于扩张稠密模块的巨大显

存消耗,AHDRNet 无法在普通的消费级显卡(如 RTX 3090 GPU, 24 GB)上推理分辨率为 2K 的 LDR 图像。

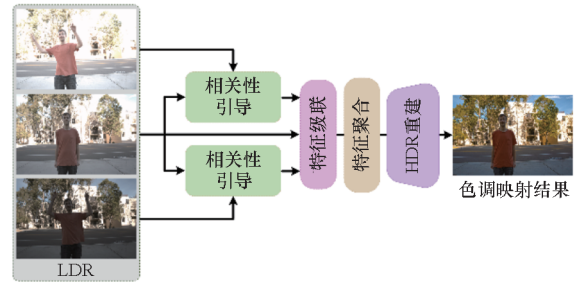


图7 基于相关性引导的处理流程

Fig. 7 Pipeline based on correlation-guided

另外一些工作在编码阶段使用特征金字塔模块来实现 LDR 特征对齐。特别地,Pu 等^[157]提出一种多尺度金字塔特征对齐策略,这种方法对处理运动和视差比较灵活和稳健。Choi 等^[158]探索 LDR 图像中像素自相似性的互注意力。为了在融合阶段以空间和通道级的方式来对齐特征,文献[158]还设计了一个双激励模块。相比之下,Yan 等^[159]不是在编码阶段而是在融合阶段应用非局部模块^[160],编码后的特征直接级联并注入融合网络。Chen 等^[161]仅使用 2 幅 LDR 图像(欠曝光和过曝光)并采用单应性网络(编码器)将欠曝光图像翘曲(warp)为过曝光图像;然后,使用注意力模块在注入融合网络之前减少非对齐的特征。在 NTIRE-HDR-2021^[93]挑战赛中,Liu 等^[162]使用金字塔、级联和可变形模块^[163]对齐 LDR 图像的特征,并用空间注意力模块来自适应地融合。

众所周知,由于 CNNs 具有局部归纳偏置的能力,其长距离建模能力依赖于网络的深度,即随着网络加深,感受野才能扩大。相比 CNNs,Transformer^[164]凭借自注意力机制来捕获全局的上下文交互关系,已经在各种视觉任务上崭露头角^[165-168]。受限于 CNNs 感受野的局部性,当面对大范围的移动和严重过饱和时,现存的基于 CNNs 多帧 HDR 的方法容易产生鬼影效应和丢失饱和区域的细节。众多结合相关性引导以及 Transformer 的去鬼影工作如雨后春笋般涌现^[46, 135, 140, 169-170]。例如,Liu 等^[81]基于 Swin Transformer^[167]和 SwinIR^[168]提出了上下文意识的视觉 Transformer(context-aware vision Transformer, CA-ViT)来重建无鬼影的 HDR 图像。Liu 等^[46]将 CA-ViT 中的空间相关性引导替换为金字塔跨尺度注意力引导,进而提出 Joint-HDRDN。考虑

到移动端 HDR 重建过程中容易受到噪声的干扰,该工作又构建了一个面向手机端的 HDR 数据集,即 Mobile-HDR。Yan 等^[169]在特征对齐阶段联合块对齐和跨注意力引导对齐,在特征聚合阶段则和上述方法保持一致,利用 Swin Transformer 模块来进行自注意力聚合。

虽然上述方法相比先前的基于卷积的方法取得了显著的性能提升,但主要存在如下问题:①特征聚合阶段的直筒式架构并未利用 Swin Transformer 的分层性优势,而该性质对上下文建模(比如遮挡以及运动物体)至关重要^[167];②预先实行特征对齐在一定程度上会在聚合的过程中累积误差,进而影响后续的特征聚合性能;③忽略了特征聚合阶段自注意力潜在的对齐作用,即上述特征对齐阶段的相关性引导属于自注意力的范畴。为此,Liu 等^[45]从分层解耦与交互聚合的视角出发,提出灵活、交互且鲁棒的(flexible, interactive and robust, FIR)框架。不同于上述的工作,FIR 引入帧间差异性估计,并将其作为提示,在逐级聚合的过程中实现“隐式对齐”,而不是预对齐,从而规避了预对齐的潜在累积误差,进而减少了运动伪影。在交互聚合的过程中,Liu 等^[45]提出了高频引导的跳跃连接策略,其加快了模型的收敛效率。此外,FIR 又提供了多种变体,其中没有特征引导的变体 FIR-U 进一步验证了自注意力的“隐式对齐”作用,即在无引导提示的情况下,该框架依旧能够在多个数据集基准上取得先进的性能。

尽管相关性引导在当前 HDR 重建框架中发挥了重要的预对齐作用,但现有基于 Transformer 的方法^[46]在聚合阶段仍难以有效填充“饱和与遮挡共存”的区域^[45]。近年来,扩散模型^[171-173]作为一类基于马尔可夫链的概率生成模型,通过可逆的噪声注入(正向过程)和去噪(反向过程)机制实现数据的生成与重构,在图像生成、图像恢复等多个视觉任务中展现了卓越的性能与广阔的应用前景。凭借其强大的生成能力^[174],文献[35, 140, 157]开始探索利用扩散模型对饱和与遮挡共存区域进行补全或幻觉生成。具体而言,GDP^[175]采用预训练的去噪扩散概率模型^[173](denoising diffusion probabilistic models, DDPM)进行图像恢复及 HDR 重建,但该方法在遮挡区域常出现填充结果不符合现实逻辑、一致性差以及推理速度较慢等问题。考虑到像素空间扩散模型计算成本高,Hu 等^[140]借鉴 Stable Diffusion^[174]在隐空间操作的思路以及 DiffIR^[176]紧凑表征的先验,

提出了一种低频感知的扩散模型 LF-Diff,通过隐空间生成低频先验使幻觉掩盖了被遮挡区域的内容结构信息。尽管在隐/潜空间进行扩散加噪与降噪能有效降低计算资源消耗,但其通常依赖于变分自编码器(variational autoencoder, VAE)^[177],而 VAE 固有的后验坍塌与表征缺陷等问题会在一定程度上削弱扩散模型的高保真重建性能。针对饱和与遮挡共存的情形,Liu 等^[45]指出利用遮挡信息而非进行虚拟补全或幻觉生成更为合理,详细讨论可参见文献[45]。

2.2.4 基于图像变换的方法

尽管基于光流的 HDR 成像能够有效实现特征对齐,但在稠密或大范围运动场景下,这类方法往往易产生较多伪影。此外,虽然基于特征级联的方法可成功重建 HDR 图像,却在过饱和区域容易出现不真实的细节。部分研究尝试在参考图像监督下,利用卷积神经网络将欠曝光与过曝光图像映射至参考帧。典型的,Rosh 等^[178]提出图像动态范围变换网络(即强度映射),将非参考帧的动态范围变换至参考帧,生成两个伪低动态范围帧;随后将参考帧与生成的两帧合并输入合成网络以重建 HDR 图像,如图 8 所示。然而,强度映射以及聚合的过程分为两阶段执行,未能实现端到端学习,导致第一阶段生成的伪曝光图像精度直接影响最终 HDR 重建结果。进一步地, Lee 等^[179]扩展了先前工作^[178],采用编解码器网络结构融合多曝光图像信息,生成精确对齐的 LDR 图像,并将对齐与合成阶段以端到端的方式联合训练。

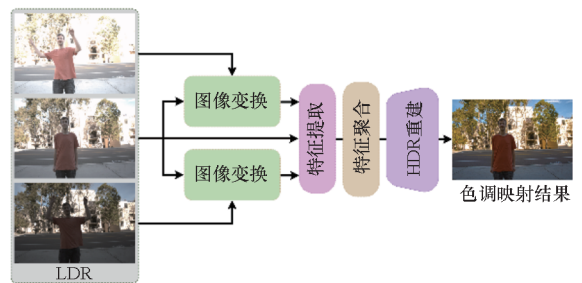


图 8 基于图像变换的处理流程

Fig. 8 Pipeline based on image translation

3 HDR 重建方法性能比较

3.1 实验细节

数据集:本文在 4 个主流的基准数据集上对最先进的方法进行了验证,包括 2 个有标签数据集,即 Kalantari 等^[54]和 Tel 等^[60];2 个没有标签的数据集,即 Sen 等^[57]和 Tursun 等^[42]。

比较方法: Sen 等^[57]、Hu 等^[121]、Kalantari 等^[54]、DeepHDR^[29]、AHDRNet^[156]、PAMNet^[157]、NHDRRnet^[159]、PSFNet^[180]、HDRRNN^[181]、HDR-GAN^[154]、APNT-Fusion^[182]、FHDRNet^[183]、STHDR^[135]、CA-ViT^[81]、SCTNet^[60]、SAFNet^[150]及其变体、FIR^[45]及其变体、DiffHDR^[35]、LFDiff^[140]和 AFUNet^[184]。

评价指标: PSNR-*l*、PSNR- μ 、SSIM-*l*、SSIM- μ 、HDR-VDP-2。

3.2 定量比较

表2是不同方法在 Kalantari 等^[54]、Tel 等^[60]

2个HDR基准数据集上的定量比较,其中部分方法因缺少代码或预训练权重,相应的指标无法计算,用“—”来标记。从表中可以看出,SAFNet及其变体、FIR及其变体、LFDiff和AFUNet在Kalantari等^[54]数据集上取得了最好的重建结果。相比SAFNet以及FIR系列,基于扩散模型的LFDiff和基于迭代优化的贝叶斯展开网络AFUNet在各项指标上较为领先。然而,在Tel等^[60]数据集上,FIR及其变体取得了最优的性能,这表明FIR分层解耦、交互聚合的范式具有较好的领域适配性及泛化性能。

表2 不同方法在两个基准数据集上的定量对比

Tab.2 Quantitative comparison of different methods on two benchmark datasets

数据集	方法	HDR-VDP-2	PSNR- <i>l</i>	SSIM- <i>l</i>	PSNR- μ	SSIM- μ
Kalantari 等 ^[54]	Sen 等 ^[57]	59.38	38.11	0.972 1	40.80	0.980 8
	Hu 等 ^[121]	—	30.76	0.950 3	35.79	0.971 7
	Kalantari 等 ^[54]	64.18	41.23	0.984 6	42.67	0.988 8
	DeepHDR ^[29]	64.90	40.88	0.985 8	41.65	0.986 0
	AHDRNet ^[156]	65.61	41.14	0.970 2	43.63	0.990 0
	PAMNet ^[157]	—	41.65	0.987 0	43.85	0.990 6
	NHDRRnet ^[159]	64.08	41.08	0.986 1	42.41	0.988 7
	PSFNet ^[180]	—	41.57	0.986 7	44.06	0.990 7
	HDRRNN ^[181]	—	41.68	—	42.07	—
	HDR-GAN ^[154]	65.54	41.57	0.986 5	43.92	0.990 5
	APNT-Fusion ^[182]	—	41.61	0.987 9	43.94	0.989 8
	FHDRNet ^[183]	65.71	41.47	0.986 9	43.91	0.990 8
	STHDR ^[135]	—	41.70	0.987 2	44.10	0.990 9
	CA-ViT ^[81]	65.66	42.18	0.988 4	44.32	0.991 6
	SCTNet ^[60]	64.48	42.29	0.988 7	44.49	0.992 4
	SAFNet-S ^[150]	66.32	42.50	0.988 8	44.11	0.991 1
	SAFNet ^[150]	<u>66.52</u>	43.18	<u>0.990 1</u>	44.66	0.991 9
	FIR ^[45]	65.92	42.50	0.989 9	44.53	0.991 8
FIR-U ^[45]	65.86	42.45	0.989 5	44.49	0.991 7	
DiffHDR ^[35]	66.67	41.73	0.988 5	44.11	0.991 1	
LFDiff ^[140]	66.38	<u>42.59</u>	0.990 6	<u>44.76</u>	0.991 9	
AFUNet ^[184]	65.92	<u>42.59</u>	0.990 6	44.91	<u>0.992 3</u>	
Tel 等 ^[60]	Kalantari 等 ^[54]	67.09	43.37	0.992 4	40.05	0.979 4
	NHDRRnet ^[159]	65.41	39.61	0.985 3	36.68	0.959 0
	AHDRNet ^[156]	68.80	45.30	0.994 3	42.08	0.983 7
	FHDRNet ^[183]	69.86	45.80	0.994 8	42.41	0.985 8
	CA-ViT ^[81]	69.23	46.35	0.994 8	42.39	0.984 4
	SCTNet ^[60]	70.66	47.51	0.995 2	42.55	0.985 0
	FIR ^[45]	71.30	48.24	0.996 1	<u>43.53</u>	<u>0.988 3</u>
	FIR-U ^[45]	<u>71.20</u>	<u>48.16</u>	0.996 1	43.62	0.988 4
	DiffHDR ^[35]	69.55	45.63	0.994 6	42.18	0.984 1
	AFUNet ^[184]	71.08	47.83	<u>0.995 9</u>	43.31	0.987 6

注: 加粗和下划线分别表示最优和次优的结果;“—”表示代码或者预训练权重没有提供,从而无法计算相应的 HDR-VDP-2。

表 3 展示了泛化性能、复杂度以及推理时间的比较结果。FIR 系列和 SAFNet 及其变体在域外泛化性能上取得了最佳的结果。值得注意的是, FIR 的卷积变体相较于基于 Transformer 的方

法,如 CA-ViT 和 SCTNet,其实现了约 60 倍的速度提升,同时保持了具有竞争力的泛化性能。这验证了 FIR 在高保真 HDR 重建与实际部署限制之间平衡的能力。

表 3 泛化性能、复杂度与推理时间的比较

Tab. 3 Comparison of generalization performance, computational complexity, and inference time

类别	方法	泛化性能				复杂度		推理时间	
		PSNR- l	SSIM- l	PSNR- μ	SSIM- μ	MACs/GB	参数量/MB	显存/GB	时间/s
CNN	AHDRNet ^[156]	43.85	0.991 4	34.16	0.961 5	725.54	<u>1.445</u>	9.069	0.348
	NHDRNet ^[159]	35.99	0.976 8	31.48	0.946 6	389.39	38.411	2.974	0.090
	FHDRNet ^[183]	43.55	0.991 4	34.53	0.962 6	1 309.44	19.534	5.643	0.327
	SAFNet-S ^[150]	<u>44.91</u>	<u>0.992 2</u>	35.27	0.9616	49.34	0.540	1.022	0.542
	SAFNet ^[150]	44.94	<u>0.992 5</u>	35.10	0.954 3	501.96	1.069	2.699	0.770
	FIR-Conv ^[45]	43.97	0.992 8	35.51	0.969 6	379.34	5.844	2.917	0.140
	FIR-Conv-Tiny ^[45]	43.87	0.992 2	35.33	<u>0.967 5</u>	214.01	3.289	<u>2.237</u>	<u>0.105</u>
	FIR-U-Conv ^[45]	44.33	0.992 8	<u>35.49</u>	<u>0.967 2</u>	378.18	5.846	2.917	0.136
Transformer	FIR-U-Conv-Tiny ^[45]	44.50	0.991 9	35.23	0.965 9	<u>213.35</u>	3.290	<u>2.237</u>	0.100
	CA-ViT ^[81]	44.65	<u>0.994 3</u>	35.86	0.973 6	609.34	<u>1.165</u>	10.637	5.590
	SCTNet ^[60]	43.37	0.994 0	35.35	0.961 8	501.96	0.948	9.732	6.621
	FIR ^[45]	44.81	0.993 7	35.31	<u>0.969 1</u>	563.32	8.163	<u>7.850</u>	2.543
	FIR-Tiny ^[45]	<u>45.05</u>	<u>0.993 9</u>	35.12	0.967 7	<u>213.42</u>	2.636	6.015	<u>0.867</u>
	FIR-U ^[45]	45.71	0.994 4	<u>35.52</u>	<u>0.967 8</u>	562.10	8.165	<u>7.850</u>	2.482
	FIR-U-Tiny ^[45]	44.59	0.993 6	35.31	0.967 6	212.73	2.637	6.015	0.773

注:所有方法均在 Kalantari 等^[54]数据集上训练后,直接在 Tel 等^[60]数据集上进行泛化性能比较;MACs 是在输入分辨率为 1 000 像素 \times 1 500 像素上计算;加粗和下划线分别表示最优和次优的结果。

3.3 定性比较

3.3.1 域外泛化测试

衡量 HDR 去鬼影算法实际应用价值的重要指标,其核心在于评估模型在训练数据分布之外的复杂场景中的适应能力。这与实验可控场景下的性能验证形成关键互补,直接决定方法能否从“算法原型”走向“实际应用”。

在 HDR 成像任务中,域外场景的差异主要源于三类核心挑战:①环境与设备的分布偏移(如训练集未覆盖的极端光照、不同相机的曝光特性差异、传感器噪声^[185]分布变化);②场景动态性的泛化需求(如复杂运动、快速光照突变、截断纹理等未见过的场景结构);③真实世界的干扰因素(如大气散射、镜头畸变、多帧对齐误差累积)。这些场景均超出模型训练时的分布范围,传统依赖训练数据拟合的方法往往会出现重影、块效应、棋盘格伪影等问题。

图 9 展示了当前最先进的方法在 Tursun

等^[42]数据集上的域外泛化性能视觉比较。显然,很多先进的方法存在明显的伪影、重影、块效应与棋盘格等问题。例如, HDRNet 中的转置卷积会引发棋盘格伪影,而 FIR 系列几乎能在复杂场景下展现了最佳的稳健性。如图 10 所示,由于 Sen 等^[57]数据集上的动态场景难度较小,几乎所有的方法均能取得较好的重建结果。

因此,域外泛化性能测试不仅通过定量指标(如表 2 中的 HDR-VDP-2、PSNR- l 等)和定性分析(伪影抑制、细节保留)衡量模型的稳健性,更能筛选出兼顾“高保真重建”与“场景适应性”的方法。例如, SAFNet 和 FIR 系列等方法,在域外测试中展现出优于传统 Transformer 基础模型的稳定性,为实际场景(如移动端实时 HDR、户外动态拍摄)的部署提供了关键支撑。

3.3.2 大范围运动

图 11 比较了不同方法在 Kalantari 等^[54]数据集上的视觉表现。当人物前景发生大范围运动时,

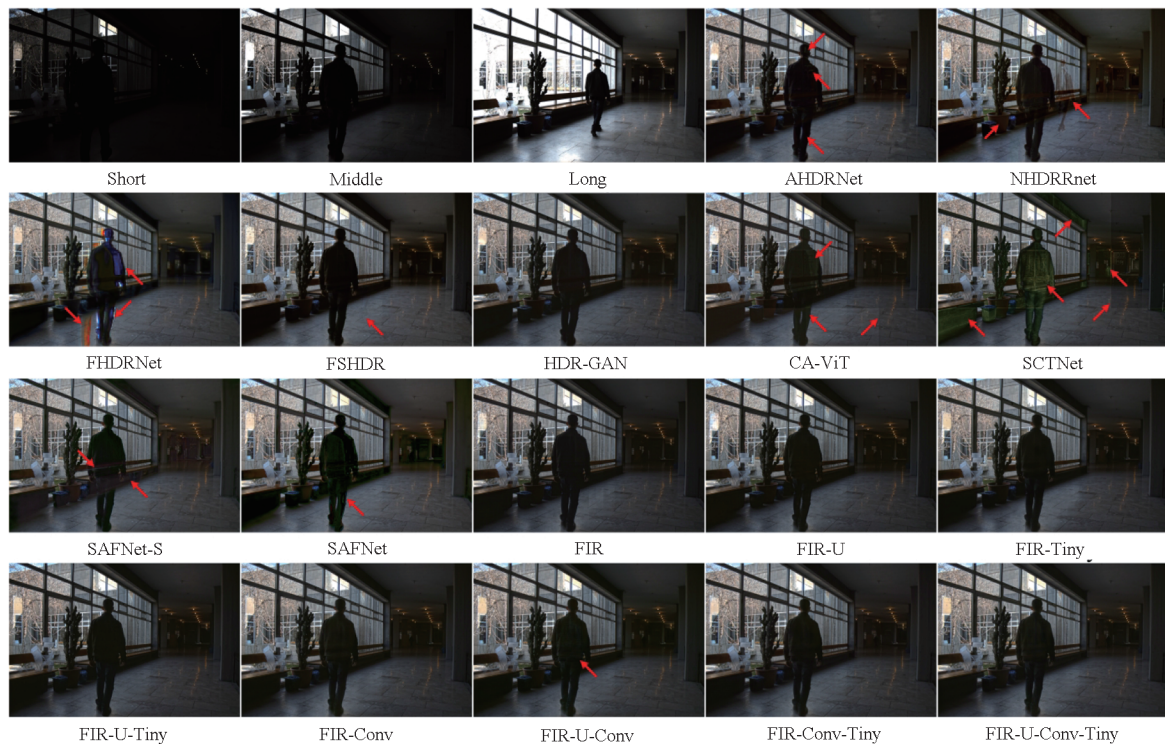


图9 不同方法在 Tursun 等数据集^[42]上的领域外视觉泛化性能对比

Fig. 9 Comparison of out-of-domain visual generalization performance of different methods on the dataset by Tursun et al. ^[42]



图10 不同方法在 Sen 等数据集^[57]上的视觉比较

Fig. 10 Visual comparison of different methods on the dataset by Sen et al. ^[57]

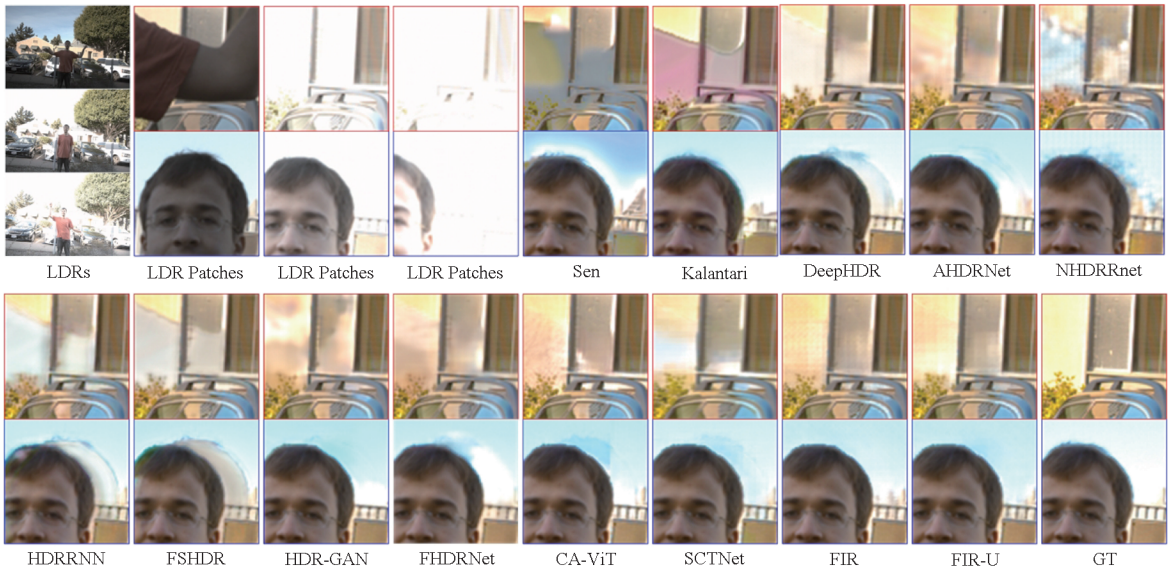


图 11 不同方法在 Kalantari 等数据集^[54]上的视觉比较

Fig. 11 Visual comparison of different methods on the dataset by Kalantari et al. ^[54]

除 FIR 及其变体外,多数方法难以实现无鬼影且高保真的重建,存在明显的运动伪影。更多关于遮挡、大范围运动的示例,请参考文献[45]及其补充材料。

4 总结与展望

本文从 HDR 定义、数据集、HDR 重建技术的发展以及实验评估四个方面对该领域进行了全面且系统的分析。综上可知,多帧 HDR 成像的核心环节在于 LDR 图像的对齐与合成。传统光流法对齐 LDR 图像往往难以保证精度,在大范围运动场景下表现尤为不佳。相比之下,基于相关性引导的特征对齐作为一种隐式对齐范式,具备更高的灵活性与有效性。因此,当前主流方法多聚焦于挖掘特征间相关性(例如通过注意力机制筛选剔除非对齐特征)。

然而,相关性引导的对齐方式对过饱和区域较为敏感,在剔除非对齐特征的过程中易造成纹理细节丢失。现有研究中,部分方法^[156]在编码阶段引入注意力机制;另有方法(如文献[45])则在解码阶段建模特征相关性,其框架的有效性可以拓展到其他底层视觉任务以及基于扩散的高分辨率生成任务中。

基于上述分析,未来 HDR 重建领域的研究可归纳为以下三个层面的关键方向:核心基础挑战、关键性能优化、前沿技术探索。

4.1 核心基础挑战

本层面聚焦于制约 HDR 重建发展的根本性问题,包括数据、去运动伪影与泛化能力。

1) 图像或特征对齐的稳健性仍是核心难题:尽管基于 LDR 图像相关性学习的对齐方案极具研究潜力,但该类方法^[156]往往伴随较高的计算开销,实用性受限。

2) 具有真实标签的 HDR 数据集的局限性:现有数据集以 Kalantari 等^[54]提出的数据集为主,其仅包含 74 组训练序列与 15 组测试序列,场景覆盖度不足,易导致模型过拟合。考虑到 HDR 拍摄设备成本较高,可通过现有先进方法生成 GT,构建多曝光 HDR 数据集。

3) 模型泛化能力提升需求迫切:为改善基于当前数据集^[54]训练模型的域外适应能力,知识迁移与半监督学习是极具前景的研究方向。

4.2 关键性能优化

本层面关注 HDR 重建技术在部署应用中的核心指标,包括:实时性、保真度与动态场景适应性。

1) 实时性优化面临技术瓶颈:深度神经网络的特征学习过程导致多数先进方法存在难以避免的推理延迟,而实时 HDR 成像是实际应用的核心需求。因此,需构建高效算法,可通过低比特量化、脉冲神经网络、脉冲 Transformer 等技术路径实现。

2) 高保真无鬼影重建技术待突破:需探索基于扩散模型^[136, 172]生成高保真、无鬼影 HDR 图像的有效方案。

3) 时序一致且高保真的 HDR 视频重建:构建帧间时序感知的端到端框架,即通过建模多帧曝光的时空相关性,解决动态场景运动模糊、

帧间闪烁、高分辨率实时推理等核心问题,适配 8K 超高清、虚拟现实/增强现实、直播等场景。当前需突破三大挑战:帧间动态信息的精准对齐与一致性维护、高分辨率数据流的轻量化处理、低带宽传输与显示终端适配的平衡。

4.3 前沿技术探索

本层面着眼于多学科交叉与新兴技术融合,探索 HDR 重建的下一代创新范式。

1) 跨模态融合与重建:例如,Event 相机与 HDR 融合^[186],依托 Event 相机高动态范围、 μs 级别的时间分辨率及低数据冗余的特性,构建“稀疏事件流-稠密 HDR 图像”的跨模态融合范式,可高效解决传统多帧 HDR 成像中动态场景运动模糊、过/欠曝光区域的细节丢失等核心问题。该融合方案在自动驾驶、移动端实时拍摄等场景中具备天然适配性,能实现低延迟、高鲁棒性的 HDR 重建。但当前仍面临事件与像素级图像的特征对齐、稀疏事件信息的稠密补充、轻量化融合模型设计等挑战,需突破跨模态表征学习与实时推理效率的平衡难题,为极端动态高动态范围场景成像提供新路径。

2) 跨模态驱动的重建范式探索:视觉语言模型^[187]驱动的 HDR 重建为该领域提供了新的研究视角,亟待进一步探索。

3) 跨任务统一框架与协同优化:构建 HDR 重建、逆色调映射、色调映射的统一框架,打破三者独立优化壁垒,以跨任务共享表征为核心,将 LDR→HDR、HDR→适配显示 LDR 全链路纳入端到端优化,通过联合损失协同实现保真度、细节恢复与显示适配的有机统一。该框架解决传统方案信息割裂、对齐误差累计问题,适配多场景需求,但需要平衡多任务目标冲突、跨场景泛化性能及轻量化部署,推动技术从分治走向协同。

4) 三维与智能化场景重建:HDR、神经辐射场^[188](neural radiance fields, NeRF)与 3D 高斯的协同技术探索,是当前高保真实时 3D 场景重建的核心升级方向。以 HDR 图像为基础数据源,提供场景中从暗部纹理到高光层次的全动态范围光影信息;借助 NeRF 完成场景结构与辐射场的建模,实现对空间中光照、材质的精准编码;再通过 3D 高斯的高效点云渲染特性,将 NeRF 构建的高保真 HDR 场景转化为可实时交互的内容,这一组合已在影视虚拟制片的实景复刻、建筑设计的全时段光影模拟等场景中展现了潜力。

5) Agent 大模型探索:通过海量 HDR/LDR 多模态图像数据学习通用视觉表征,能够赋予模型跨场景 HDR 重建迁移能力。例如,适配室内外逆光、户外强光、大光比等差异化场景;在产业化落地场景中,该类模型展现了高效的统一参数优化潜力,但需在模型规模与低功耗实时处理的实际需求之间寻求平衡。

参考文献 (References)

- [1] MANTIUK R K, MYSZKOWSKI K, SEIDEL H P. High dynamic range imaging [M]//WEBSTER J G. Wiley Encyclopedia of Electrical and Electronics Engineering. Hoboken: John Wiley & Sons, Ltd, 2015: 1-42.
- [2] REINHARD E. High dynamic range imaging[M]//IKEUCHI K. Computer Vision. Cham: Springer, 2021: 558-563.
- [3] BANTERLE F, ARTUSI A, DEBATTISTA K, et al. Advanced high dynamic range imaging [M]. 2nd ed. Boca Raton: A K Peters/CRC Press, 2017.
- [4] LIN Y T, CHEN S K, HU H N, et al. HDR reconstruction boosting with training-free and exposure-consistent diffusion[EB/OL]. (2026-02-23) [2025-10-25]. <https://arxiv.org/abs/2602.19706>.
- [5] CANHAM T D, TEDLA S, MURDOCH M J, et al. Gain-MLP: improving HDR gain map encoding via a lightweight MLP[EB/OL]. (2025-03-14) [2025-10-25]. <https://arxiv.org/abs/2503.11883>.
- [6] DEBEVEC P E, MALIK J. Recovering high dynamic range radiance maps from photographs[C]//Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, 1997: 369-378.
- [7] SRIKANTHA A, SIDIBÉ D. Ghost detection and removal for high dynamic range images: recent advances [J]. Signal Processing: Image Communication, 2012, 27(6): 650-662.
- [8] CAO P B, CHEN H Y, MA J Z, et al. Learned HDR image compression for perceptually optimal storage and display[C]// Proceedings of European Conference on Computer Vision, 2024: 109-126.
- [9] SHU Y, SHEN L Q, HU X Y, et al. Towards real-world HDR video reconstruction: a large-scale benchmark dataset and a two-stage alignment network[C]//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 2879-2888.
- [10] MA K D, ZENG K, WANG Z. Perceptual quality assessment for multi-exposure image fusion [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3345-3356.
- [11] MERTENS T, KAUTZ J, VAN REETH F. Exposure fusion[C]//Proceedings of the 15th Pacific Conference on Computer Graphics and Applications (PG'07), 2007: 382-390.
- [12] VELICHKO S, JASINSKI D, GUIDASH M, et al. Automotive 3 μm HDR image sensor with LFM and distance functionality [J]. IEEE Transactions on Electron Devices, 2022, 69(6): 2951-2956.
- [13] SUONSIVU A, SALMELA L, HELIN L, et al. Color preserving CMOS-SPAD fusion for multi-frame HDR [EB/OL]. [2025-10-25]. <https://openaccess.thecvf.com/>

- content/WACV2026/papers/Suonsivu_Color_Preserving_CMOS-SPAD_Fusion_for_Multi-Frame_HDR_WACV_2026_paper.pdf.
- [14] COGALAN U, BEMANA M, MYSZKOWSKI K, et al. Learning HDR video reconstruction for dual-exposure sensors with temporally-alternating exposures [J]. *Computers & Graphics*, 2022, 105: 57–72.
- [15] SIDDIQ A, AHMED J, KHAN I R. GiTMO: an HDR10 compatible generic tone-mapping for real-time video play[J]. *IEEE Transactions on Consumer Electronics*, 2022, 68(4): 411–418.
- [16] CAO C, YUE H J, LIU X, et al. Unsupervised HDR image and video tone mapping via contrastive learning[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(2): 786–798.
- [17] CAO P B, LE C Y, FANG Y M, et al. A perceptually optimized and self-calibrated tone mapping operator [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2025, 31(10): 8268–8282.
- [18] ENDO Y, KANAMORI Y, MITANI J. Deep reverse tone mapping[J]. *ACM Transactions on Graphics*, 2017, 36(6): 177.
- [19] LE P H, LE Q, NGUYEN R, et al. Single-image HDR reconstruction by multi-exposure generation[C]//*Proceedings of 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2023: 4052–4061.
- [20] WANG C, SERRANO A, PAN X G, et al. GlowGAN: unsupervised learning of HDR images from LDR images in the wild [C]//*Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023: 10475–10485.
- [21] BANITALEBI-DEHKORDI A, AZIMI M, POURAZAD M T, et al. Compression of high dynamic range video using the HEVC and H. 264/AVC standards [C]//*Proceedings of the 10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, 2014: 8–12.
- [22] MANTIUK R, EFREMOV A, MYSZKOWSKI K, et al. Backward compatible high dynamic range MPEG video compression [J]. *ACM Transactions on Graphics*, 2006, 25(3): 713–723.
- [23] WANG H, YE M, ZHU X T, et al. Compressed-SDR to HDR video reconstruction[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(5): 3679–3691.
- [24] AZIMI M, BOITARD R, POURAZAD M T, et al. Performance evaluation of single layer HDR video transmission pipelines[J]. *IEEE Transactions on Consumer Electronics*, 2017, 63(3): 267–276.
- [25] GUO C, FAN L D, XUE Z Y, et al. Learning a practical SDR-to-HDRTV up-conversion using new dataset and degradation models [C]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023: 22231–22241.
- [26] TURSUN O T, AKYÜZ A O, ERDEM A, et al. An objective deghosting quality metric for HDR images [J]. *Computer Graphics Forum*, 2016, 35(2): 139–152.
- [27] MANTIUK R, KIM K J, REMPEL A G, et al. HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions [J]. *ACM Transactions on Graphics*, 2011, 30(4): 1–14.
- [28] KARADUZOVIC-HADZIABDIC K, TELALOVIC J H, MANTIUK R. Subjective and objective evaluation of multi-exposure high dynamic range image deghosting methods [C]//*Proceedings of the 37th Annual Conference of the European Association for Computer Graphics: Short Papers*, 2016: 29–32.
- [29] WU S Z, XU J R, TAI Y W, et al. Deep high dynamic range imaging with large foreground motions [C]//*Proceedings of European Conference on Computer Vision*, 2018: 120–135.
- [30] XU G W, WANG Y J, GU J W, et al. HDRFlow: real-time HDR video reconstruction with large motions [C]//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024: 24851–24860.
- [31] CHUNG H, CHO N I. High dynamic range imaging of dynamic scenes with saturation compensation but without explicit motion compensation [C]//*Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2022: 61–71.
- [32] CACCIOTTI M, CAMUS V, SCHLACHTER J, et al. Hardware acceleration of HDR-image tone mapping on an FPGA-CPU platform through high-level synthesis [C]//*Proceedings of the 31st IEEE International System-on-Chip Conference (SOCC)*, 2018: 158–162.
- [33] NI Z K, ZHANG Y, REN K R, et al. Semantic masking with curriculum learning for robust HDR image reconstruction[J]. *International Journal of Computer Vision*, 2025, 133(10): 6896–6911.
- [34] VIEN A G, LEE C. Exposure-aware dynamic weighted learning for single-shot HDR imaging [C]//*Proceedings of the 17th European Conference on Computer Vision*, 2022: 435–452.
- [35] YAN Q S, HU T, SUN Y, et al. Toward high-quality HDR deghosting with conditional diffusion models [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(5): 4011–4026.
- [36] WANG L, YOON K J. Deep learning for HDR imaging: state-of-the-art and future trends [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(12): 8874–8895.
- [37] WANG H, YE M, LUO D Y, et al. Multi-exposure high dynamic range reconstruction by incorporating imaging knowledge [J]. *Engineering Applications of Artificial Intelligence*, 2026, 169: 114177.
- [38] QI Y S, ZHU L, BAO N, et al. Seeing through light and darkness: sensor-physics grounded deblurring HDR NeRF from single-exposure images and events [EB/OL]. (2026–01–21) [2026–03–22]. <https://arxiv.org/abs/2601.15475>.
- [39] BANTERLE F, MARNERIDES D, BASHFORD-ROGERS T, et al. Self-supervised high dynamic range imaging: what can be learned from a single 8-bit video [J]. *ACM Transactions on Graphics*, 2024, 43(2): 1–16.
- [40] WANG H, ZHANG C B, QIAN X, et al. HDRSL net for accurate high dynamic range imaging-based structured light 3D reconstruction [J]. *IEEE Transactions on Image Processing*, 2025, 34: 5486–5499.
- [41] YE Y Y, ZHANG N, ZHAO Y, et al. Deep video inverse tone mapping based on temporal clues [C]//*Proceedings of*

- 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 25995 – 26004.
- [42] TURSUN O T, AKYÜZ A O, ERDEM A, et al. The state of the art in HDR deghosting: a survey and evaluation [J]. *Computer Graphics Forum*, 2015, 34(2): 683 – 707.
- [43] TOPIWALA P, DAI W, KRISHNAN M. Improvements on HDR10[C]//Proceedings of 2016 Digital Media Industry & Academic Forum (DMIAF), 2016: 17 – 22.
- [44] JUN-SEONG K, YU-JI K, YE-BIN M, et al. HDR-pixeloxels: self-calibrating high dynamic range radiance fields[C]//Proceedings of European Conference on Computer Vision, 2022: 384 – 401.
- [45] LIU X N, LI A, WU Z W, et al. PASTA: towards flexible and efficient HDR imaging via progressively aggregated spatio-temporal alignment [EB/OL]. (2024 – 04 – 09) [2025 – 10 – 25]. <https://arxiv.org/abs/2403.10376>.
- [46] LIU S Z, ZHANG X D, SUN L C, et al. Joint HDR denoising and fusion: a real-world mobile HDR image dataset[C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 13966 – 13975.
- [47] HU L T, CHEN H J, ALLEBACH J P. Joint multi-scale tone mapping and denoising for HDR image enhancement [C]//Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW), 2022: 729 – 738.
- [48] CHEN X Y, LIU Y H, ZHANG Z W, et al. HDRUNet: single image HDR reconstruction with denoising and dequantization [C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2021: 354 – 363.
- [49] WEN Q, RAO Z F, LEI C Y, et al. Enhancing HDR imaging with joint denoising and deblurring[J]. *International Journal of Computer Vision*, 2025, 133(11): 7536 – 7552.
- [50] SEN P, AGUERREBERE C. Practical high dynamic range imaging of everyday scenes: photographing the world as we see it with our own eyes[J]. *IEEE Signal Processing Magazine*, 2016, 33(5): 36 – 44.
- [51] MANN S. Comparametric equations with practical applications in quantigraphic image processing[J]. *IEEE Transactions on Image Processing*, 2000, 9(8): 1389 – 1406.
- [52] HORN B K P. *Robot vision*[M]. Cambridge: MIT Press, 1986.
- [53] IMPOCO G, UKOVICH A. Exploiting the local contrast to evaluate image dynamics reduction algorithms [EB/OL]. [2025 – 10 – 25]. <https://ricerca.unifyvg.it/entities/publication/b1333049-b39957410c5b/details>.
- [54] KALANTARI N K, RAMAMOORTHY R. Deep high dynamic range imaging of dynamic scenes[J]. *ACM Transactions on Graphics*, 2017, 36(4): 1 – 12.
- [55] HASINOFF S W, DURAND F, FREEMAN W T. Noise-optimal capture for high dynamic range photography [C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010: 553 – 560.
- [56] GRANADOS M, AJDIN B, WAND M, et al. Optimal HDR reconstruction with linear digital cameras[C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010: 215 – 222.
- [57] SEN P, KALANTARI N K, YAESOUBI M, et al. Robust patch-based HDR reconstruction of dynamic scenes[J]. *ACM Transactions on Graphics*, 2012, 31(6): 203.
- [58] PRABHAKAR K R, ARORA R, SWAMINATHAN A, et al. A fast, scalable, and reliable deghosting method for extreme exposure fusion[C]//Proceedings of 2019 IEEE International Conference on Computational Photography (ICCP), 2019: 1 – 8.
- [59] HU J H, CHOE G, NADIR Z, et al. Sensor-realistic synthetic data engine for multi-frame high dynamic range photography [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020: 2180 – 2189.
- [60] TEL S, WU Z W, ZHANG Y L, et al. Alignment-free HDR deghosting with semantics consistent transformer [EB/OL]. (2023 – 09 – 28) [2025 – 10 – 25]. <https://arxiv.org/abs/2305.18135>.
- [61] EILERTSEN G, KRONANDER J, DENES G, et al. HDR image reconstruction from a single exposure using deep CNNs[J]. *ACM Transactions on Graphics (TOG)*, 2017, 36(6): 1 – 15.
- [62] LIU Y L, LAI W S, CHEN Y S, et al. Single-image HDR reconstruction by learning to reverse the camera pipeline[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 1648 – 1657.
- [63] ZHANG J S, LALONDE J F. Learning high dynamic range from outdoor panoramas [C]//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV), 2017: 4529 – 4538.
- [64] NEMOTO H, KORSHUNOV P, HANHART P, et al. Visual attention in LDR and HDR images[C]//Proceedings of the 9th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), 2015.
- [65] KIM S Y, OH J, KIM M. Deep SR-ITM: joint learning of super-resolution and inverse tone-mapping for 4K UHD HDR applications [C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 3116 – 3125.
- [66] CHEN G Y, CHEN C F, GUO S, et al. HDR video reconstruction: a coarse-to-fine network and a real-world benchmark dataset [C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 2482 – 2491.
- [67] KALANTARI N K, SHECHTMAN E, BARNES C, et al. Patch-based high dynamic range video[J]. *ACM Transactions on Graphics*, 2013, 32(6): 1 – 8.
- [68] MA K D, YEGANEH H, ZENG K, et al. High dynamic range image compression by optimizing tone mapped image quality index[J]. *IEEE Transactions on Image Processing*, 2015, 24(10): 3086 – 3097.
- [69] CAI J R, GU S H, ZHANG L, et al. Learning a deep single image contrast enhancer from multi-exposure images [J]. *IEEE Transactions on Image Processing*, 2018, 27(4): 2049 – 2062.
- [70] LEE S, AN G H, KANG S J. Deep chain HDRI: reconstructing a high dynamic range image from a single low dynamic range image[J]. *IEEE Access*, 2018, 6: 49913 – 49924.
- [71] ZOLL C. Optimal tone mapping [J]. *Linguistic Inquiry*, 2003, 34(2): 225 – 268.

- [72] DURAND F, DORSEY J. Interactive tone mapping [C]// Proceedings of Eurographics Workshop on Rendering Techniques, 2000: 219–230.
- [73] SALIH Y, MD-ESA W, MALIK A S, et al. Tone mapping of HDR images: a review [C]// Proceedings of the 4th International Conference on Intelligent and Advanced Systems (ICIAS2012), 2012: 368–373.
- [74] EILERTSEN G, MANTIUK R K, UNGER J. A comparative review of tone-mapping algorithms for high dynamic range video[J]. Computer Graphics Forum, 2017, 36(2): 565–592.
- [75] MANTIUK R, DALY S, KEROFISKY L. Display adaptive tone mapping [C]// Proceedings of Special Interest Group on Computer Graphics and Interactive Techniques Conference, 2008: 68.
- [76] BORREGO-CARAZO J, OZAY M, LABOYRIE F, et al. A mixed quantization network for computationally efficient mobile inverse tone mapping [EB/OL]. (2022–03–12) [2025–10–25]. <https://arxiv.org/abs/2203.06504>.
- [77] GUARNIERI G, MARSÌ S, RAMPONI G. High dynamic range image display with halo and clipping prevention [J]. IEEE Transactions on Image Processing, 2011, 20(5): 1351–1362.
- [78] REMPEL A G, TRENTACOSTE M, SEETZEN H, et al. Ldr2Hdr: on-the-fly reverse tone mapping of legacy video and photographs [J]. ACM Transactions on Graphics (TOG), 2007, 26(3): 39.
- [79] MASIA B, AGUSTIN S, FLEMING R W, et al. Evaluation of reverse tone mapping through varying exposure conditions [J]. ACM Transactions on Graphics, 2009, 28(5): 1–8.
- [80] KOVALESKI R P, OLIVEIRA M M. High-quality reverse tone mapping for a wide range of exposures [C]// Proceedings of the 27th SIBGRAPI Conference on Graphics, Patterns and Images, 2014: 49–56.
- [81] LIU Z, WANG Y L, ZENG B, et al. Ghost-free high dynamic range imaging with context-aware transformer [C]// Proceedings of European Conference on Computer Vision, 2022: 344–360.
- [82] ALPAY K C, AKYÜZ A O, BRANDONISIO N, et al. DeepDuoHDR: a low complexity two exposure algorithm for HDR deghosting on mobile devices [J]. IEEE Transactions on Image Processing, 2024, 33: 6592–6606.
- [83] WANG Z, BOVIK A C. A universal image quality index [J]. IEEE Signal Processing Letters, 2002, 9(3): 81–84.
- [84] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600–612.
- [85] MANTIUK R, DALY S J, MYSZKOWSKI K, et al. Predicting visible differences in high dynamic range images: model and its calibration [C]// Proceedings of Human Vision and Electronic Imaging X, 2005: 204–214.
- [86] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment [C]// Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003: 1398–1402.
- [87] NARWARIA M, DA SILVA M P, LE CALLET P. HDR-VQM: an objective quality measure for high dynamic range video [J]. Signal Processing: Image Communication, 2015, 35: 46–60.
- [88] WANG J, FAN H T, HOU X X, et al. MSTRIQ: no reference image quality assessment based on swin transformer with multi-stage fusion [C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022: 1268–1277.
- [89] YANG S D, WU T H, SHI S W, et al. MANIQA: multi-dimension attention network for no-reference image quality assessment [C]// Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022: 1190–1199.
- [90] FANG Y M, ZHU H W, ZENG Y, et al. Perceptual quality assessment of smartphone photography [C]// Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 3674–3683.
- [91] JIA S, ZHANG Y, AGRAFIOTIS D, et al. Blind high dynamic range image quality assessment using deep learning [C]// Proceedings of 2017 IEEE International Conference on Image Processing (ICIP), 2017: 765–769.
- [92] HE Q, LI D Q, JIANG T T, et al. Quality assessment for tone-mapped HDR images using multi-scale and multi-layer information [C]// Proceedings of 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2018: 1–6.
- [93] PÉREZ-PELLITERO E, CATLEY-CHANDAR S, LEONARDIS A, et al. NTIRE 2021 challenge on high dynamic range imaging: dataset, methods and results [C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2021.
- [94] PÉREZ-PELLITERO E, CATLEY-CHANDAR S, SHAW R, et al. NTIRE 2022 challenge on high dynamic range imaging: methods and results [C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2022: 1009–1023.
- [95] LEE S, PARK E, CANELO A, et al. NTIRE 2025 challenge on efficient burst HDR and restoration: datasets, methods, and results [C]// Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2025: 993–1008.
- [96] FROELICH J, GRANDINETTI S, EBERHARDT B, et al. Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays [C]// Proceedings of IS&T/SPIE Electronic Imaging, 2014: 90230X.
- [97] KHAN E A, AKYUZ A O, REINHARD E. Ghost removal in high dynamic range images [C]// Proceedings of International Conference on Image Processing, 2006: 2005–2008.
- [98] JACOBS K, LOSCOS C, WARD G. Automatic high-dynamic range image generation for dynamic scenes [J]. IEEE Computer Graphics and Applications, 2008, 28(2): 84–93.
- [99] PECE F, KAUTZ J. Bitmap movement detection: HDR for dynamic scenes [C]// Proceedings of Conference on Visual Media Production, 2010: 1–8.
- [100] ZHANG W, CHAM W K. Gradient-directed multiexposure composition [J]. IEEE Transactions on Image Processing, 2012, 21(4): 2318–2323.
- [101] GROSCHE T. Fast and robust high dynamic range image generation with camera and object movement [EB/OL]. [2025–10–25]. <https://www2.in.tu-clausthal.de/~>

- cgstore/publications/old/robust_hdr_vmv06.pdf.
- [102] GALLO O, GELFANDZ N, CHEN W C, et al. Artifact-free high dynamic range imaging [C]//Proceedings of IEEE International Conference on Computational Photography (ICCP), 2009; 1 - 7.
- [103] RAMAN S, CHAUDHURI S. Reconstruction of high contrast images for dynamic scenes [J]. *The Visual Computer*, 2011, 27(12): 1099 - 1114.
- [104] HEO Y S, LEE K M, LEE S U, et al. Ghost-free high dynamic range imaging [C]//Proceedings of Asian Conference on Computer Vision, 2010; 486 - 500.
- [105] GRANADOS M, KIM K I, TOMPKIN J, et al. Automatic noise modeling for ghost-free HDR reconstruction [J]. *ACM Transactions on Graphics*, 2013, 32(6): 1 - 10.
- [106] ZHENG J H, LI Z G, ZHU Z J, et al. Hybrid patching for a sequence of differently exposed images with moving objects [J]. *IEEE Transactions on Image Processing*, 2013, 22(12): 5190 - 5201.
- [107] LI Z G, ZHENG J H, ZHU Z J, et al. Selectively detail-enhanced fusion of differently exposed images with moving objects [J]. *IEEE Transactions on Image Processing*, 2014, 23(10): 4372 - 4382.
- [108] OH T H, LEE J Y, TAI Y W, et al. Robust high dynamic range imaging by rank minimization [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(6): 1219 - 1232.
- [109] LEE C, LI Y L, MONGA V. Ghost-free high dynamic range imaging via rank minimization [J]. *IEEE Signal Processing Letters*, 2014, 21(9): 1045 - 1049.
- [110] WARD G. Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures [J]. *Journal of Graphics Tools*, 2003, 8(2): 17 - 30.
- [111] LIU Z W, YUAN L, TANG X O, et al. Fast burst images denoising [J]. *ACM Transactions on Graphics*, 2014, 33(6): 1 - 9.
- [112] AKYÜZ A O. Photographically guided alignment for HDR images [C]//Proceedings of Eurographics, 2011; 73 - 74.
- [113] YAO S. Robust image registration for multiple exposure high dynamic range image synthesis [C]//Proceedings of Image Processing: Algorithms and Systems IX, 2011; 78700Q.
- [114] GALLO O, TROCCOLI A, HU J, et al. Locally non-rigid registration for mobile HDR photography [C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015; 48 - 55.
- [115] HU J, GALLO O, PULLI K. Exposure stacks of live scenes with hand-held cameras [C]//Proceedings of European Conference on Computer Vision, 2012; 499 - 512.
- [116] KANG S B, UYTENDAELE M, WINDER S, et al. High dynamic range video [J]. *ACM Transactions On Graphics*, 2003, 22(3): 319 - 325.
- [117] BOGONI L. Extending dynamic range of monochrome and color images through fusion [C]//Proceedings of the 15th International Conference on Pattern Recognition, 2000; 7 - 12.
- [118] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision [C]//Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981; 674 - 679.
- [119] JINNO T, OKUDA M. Motion blur free HDR image acquisition using multiple exposures [C]//Proceedings of the 15th IEEE International Conference on Image Processing, 2008; 1304 - 1307.
- [120] ZIMMER H, BRUHN A, WEICKERT J. Freehand HDR imaging of moving scenes with simultaneous resolution enhancement [J]. *Computer Graphics Forum*, 2011, 30(2): 405 - 414.
- [121] HU J, GALLO O, PULLI K, et al. HDR deghosting: how to deal with saturation? [C]//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013; 1163 - 1170.
- [122] HACOHEN Y, SHECHTMAN E, GOLDMAN D B, et al. Non-rigid dense correspondence with applications for image enhancement [J]. *ACM Transactions on Graphics*, 2011, 30(4): 1 - 10.
- [123] FATTAL R, LISCHINSKI D, WERMAN M. Gradient domain high dynamic range compression [M]//Seminal Graphics Papers: Pushing the Boundaries, Volume 2. New York: Association for Computing Machinery, 2023; 671 - 678.
- [124] EDEN A, UYTENDAELE M, SZELISKI R. Seamless image stitching of scenes with large motions and exposure differences [C]//Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006; 2498 - 2505.
- [125] LU P Y, HUANG T H, WU M S, et al. High dynamic range image reconstruction from hand-held cameras [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2009; 509 - 516.
- [126] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. PatchMatch; a randomized correspondence algorithm for structural image editing [J]. *ACM Transactions on Graphics*, 2009, 28(3): 1 - 11.
- [127] ZHENG J H, LI Z G, ZHU Z J, et al. Patching of moving objects for ghosting-free HDR synthesis [C]//Proceedings of SIGGRAPH '12; ACM SIGGRAPH 2012 Posters, 2012; 62.
- [128] MA K D, LI H, YONG H W, et al. Robust multi-exposure image fusion; a structural patch decomposition approach [J]. *IEEE Transactions on Image Processing*, 2017, 26(5): 2519 - 2532.
- [129] MA K D, WANG Z. Multi-exposure image fusion: a patch-wise approach [C]//Proceedings of 2015 IEEE International Conference on Image Processing, 2015; 1717 - 1721.
- [130] GROSSBERG M D, NAYAR S K. Determining the camera response from images; what is knowable? [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(11): 1455 - 1467.
- [131] LI H, MA K D, YONG H W, et al. Fast multi-scale structural patch decomposition for multi-exposure image fusion [J]. *IEEE Transactions on Image Processing*, 2020, 29: 5805 - 5816.
- [132] LI H, CHAN T N, QI X B, et al. Detail-preserving multi-exposure fusion with edge-preserving structural patch decomposition [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(11): 4293 - 4304.
- [133] PRABHAKAR K R, SENTHIL G, AGRAWAL S, et al. Labeled from unlabeled; exploiting unlabeled data for few-shot deep HDR deghosting [C]//Proceedings of IEEE/CVF

- Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 4873–4883.
- [134] YAN Q S, ZHANG S, CHEN W Y, et al. SMAE: few-shot learning for HDR deghosting with saturation-aware masked autoencoders [C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 5775–5784.
- [135] SONG J W, PARK Y I, KONG K, et al. Selective TransHDR: transformer-based selective HDR imaging using ghost region mask [C]//Proceedings of European Conference On Computer Vision, 2022: 288–304.
- [136] GUAN Y S, XU R K, YAO M D, et al. Diffusion-promoted HDR video reconstruction [C]//Proceedings of European Conference on Computer Vision, 2024: 20–38.
- [137] DALAL D, VASHISHTHA G, SINGH P, et al. Single image LDR to HDR conversion using conditional diffusion [C]//Proceedings of IEEE International Conference on Image Processing (ICIP), 2023: 3533–3537.
- [138] WANG C, XIA Z H, LEIMKÜHLER T, et al. LEDiff: latent exposure diffusion for HDR generation [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025: 453–464.
- [139] BEMANA M, LEIMKÜHLER T, MYŚKOWSKI K, et al. Bracket diffusion: HDR image generation by consistent LDR denoising [J]. Computer Graphics Forum, 2025, 44(2): e70086.
- [140] HU T, YAN Q S, QI Y K, et al. Generating content for HDR deghosting from frequency view [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 25732–25741.
- [141] CHEN Z X, WANG Y J, CAI X, et al. UltraFusion: ultra high dynamic imaging using exposure fusion [C]//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025: 16111–16121.
- [142] LIU C, FREEMAN W T, ADELSON E H, et al. Human-assisted motion annotation [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2008: 1–8.
- [143] LIU C. Beyond pixels: exploring new representations and applications for motion analysis [D]. Cambridge: Massachusetts Institute of Technology, 2009.
- [144] LIU C, YUEN J, TORRALBA A. SIFT flow: dense correspondence across scenes and its applications [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 978–994.
- [145] YAN Q S, GONG D, ZHANG P P, et al. Multi-scale dense networks for deep high dynamic range imaging [C]//Proceedings of 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), 2019: 41–50.
- [146] PENG F Y, ZHANG M J, LAI S M, et al. Deep HDR reconstruction of dynamic scenes [C]//Proceedings of IEEE the 3rd International Conference on Image, Vision and Computing (ICIVC), 2018: 347–351.
- [147] DOSOVITSKIY A, FISCHER P, ILG E, et al. FlowNet: learning optical flow with convolutional networks [C]//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 2758–2766.
- [148] ILG E, MAYER N, SAIKIA T, et al. FlowNet 2.0: evolution of optical flow estimation with deep networks [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1647–1655.
- [149] CATLEY-CHANDAR S, TANAY T, VANDROUX L, et al. FlexHDR: modeling alignment and exposure uncertainties for flexible HDR imaging [J]. IEEE Transactions on Image Processing, 2022, 31: 5923–5935.
- [150] KONG L T, LI B, XIONG Y K, et al. SAFNet: selective alignment fusion network for efficient HDR imaging [C]//Proceedings of European Conference on Computer Vision, 2024: 256–273.
- [151] YAN Q S, WANG B, ZHANG L, et al. Towards accurate HDR imaging with learning generator constraints [J]. Neurocomputing, 2021, 428: 79–91.
- [152] CHAUDHARI P, SCHIRRMACHER F, MAIER A, et al. Merging-ISP: multi-exposure high dynamic range image signal processing [C]//Proceedings of DAGM German Conference on Pattern Recognition, 2021: 328–342.
- [153] METWALY K, MONGA V. Attention-mask dense merger (attendense) deep HDR for ghost removal [C]//Proceedings of ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020: 2623–2627.
- [154] NIU Y Z, WU J B, LIU W X, et al. HDR-GAN: HDR image reconstruction from multi-exposed LDR images with large motions [J]. IEEE Transactions on Image Processing, 2021, 30: 3885–3896.
- [155] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks [C]//Proceedings of the 29th International Conference on Neural Information Processing Systems, 2015: 2017–2025.
- [156] YAN Q S, GONG D, SHI Q F, et al. Attention-guided network for ghost-free high dynamic range imaging [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 1751–1760.
- [157] PU Z Y, GUO P Y, ASIF M S, et al. Robust high dynamic range (HDR) imaging with complex motion and parallax [C]//Proceedings of Asian Conference on Computer Vision, 2020: 134–149.
- [158] CHOI S, CHO J, SONG W, et al. Pyramid inter-attention for high dynamic range imaging [J]. Sensors, 2020, 20(18): 5102.
- [159] YAN Q S, ZHANG L, LIU Y, et al. Deep HDR imaging via a non-local network [J]. IEEE Transactions on Image Processing, 2020, 29: 4308–4322.
- [160] WANG X L, GIRSHICK R, GUPTA A, et al. Non-local neural networks [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7794–7803.
- [161] CHEN S Y, CHUANG Y Y. Deep exposure fusion with deghosting via homography estimation and attention learning [C]//Proceedings of ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020: 1464–1468.
- [162] LIU Z, LIN W J, LI X P, et al. ADNet: attention-guided deformable convolutional network for high dynamic range imaging [C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2021: 463–470.
- [163] WANG X T, CHAN K C K, YU K, et al. EDVR: video restoration with enhanced deformable convolutional

- networks[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019: 1954 – 1963.
- [164] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000 – 6010.
- [165] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16 × 16 words; transformers for image recognition at scale[EB/OL]. (2021 – 06 – 03) [2025 – 10 – 25]. <https://arxiv.org/abs/2010.11929>.
- [166] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]//Proceedings of European Conference on Computer Vision, 2020: 213 – 229.
- [167] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 9992 – 10002.
- [168] LIANG J Y, CAO J Z, SUN G L, et al. SwinIR: image restoration using swin transformer [C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2021: 1833 – 1844.
- [169] YAN Q S, CHEN W Y, ZHANG S, et al. A unified HDR imaging method with pixel and patch level [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 22211 – 22220.
- [170] CHEN R F, ZHENG B L, ZHANG H, et al. Improving dynamic HDR imaging with fusion transformer [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(1): 340 – 349.
- [171] SONG Y, SOHL-DICKSTEIN J, KINGMA D P, et al. Score-based generative modeling through stochastic differential equations[EB/OL]. (2021 – 02 – 10) [2025 – 10 – 25]. <https://arxiv.org/abs/2011.13456>.
- [172] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[EB/OL]. (2000 – 12 – 16) [2025 – 10 – 25]. <https://arxiv.org/abs/2006.11239>.
- [173] NICHOL A Q, DHARIWAL P. Improved denoising diffusion probabilistic models [C]//Proceedings of the 38th International Conference on Machine Learning, 2021: 8162 – 8171.
- [174] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 10674 – 10685.
- [175] FEI B, LYU Z Y, PAN L, et al. Generative diffusion prior for unified image restoration and enhancement [C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 9935 – 9946.
- [176] XIA B, ZHANG Y L, WANG S Y, et al. DiffIR: efficient diffusion model for image restoration [C]//Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 13049 – 13059.
- [177] KINGMA D P, WELING M. Auto-encoding variational bayes[EB/OL]. (2022 – 12 – 10) [2025 – 10 – 25]. <https://arxiv.org/abs/1312.6114>.
- [178] K S G R, BISWAS A, PATEL M S, et al. Deep multi-stage learning for HDR with large object motions [C]//Proceedings of IEEE International Conference on Image Processing (ICIP), 2019: 4714 – 4718.
- [179] LEE S H, CHUNG H, CHO N I. Exposure-structure blending network for high dynamic range imaging of dynamic scenes [J]. IEEE Access, 2020, 8: 117428 – 117438.
- [180] YE Q, XIAO J, LAM K M, et al. Progressive and selective fusion network for high dynamic range imaging [C]//Proceedings of the 29th ACM International Conference on Multimedia, 2021: 5290 – 5297.
- [181] PRABHAKAR K R, AGRAWAL S, BABU R V. Self-gated memory recurrent network for efficient scalable HDR deghosting [J]. IEEE Transactions on Computational Imaging, 2021, 7: 1228 – 1239.
- [182] CHEN J, YANG Z F, CHAN T N, et al. Attention-guided progressive neural texture fusion for high dynamic range image restoration [J]. IEEE Transactions on Image Processing, 2022, 31: 2661 – 2672.
- [183] DAI T H, LI W, CAO X L, et al. Wavelet-based network for high dynamic range imaging [J]. Computer Vision and Image Understanding, 2024, 238: 103881.
- [184] LI X Y, NI Z K, YANG W H. AFUNet: cross-iterative alignment-fusion synergy for HDR reconstruction via deep unfolding paradigm [EB/OL]. (2025 – 07 – 05) [2025 – 10 – 25]. <https://arxiv.org/abs/2506.23537>.
- [185] CHI Y H, ZHANG X G, CHAN S H. HDR imaging with spatially varying signal-to-noise ratios [C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 5724 – 5734.
- [186] MOSTAFAVI M, WANG L, YOON K J. Learning to reconstruct HDR images from events, with applications to depth and flow prediction [J]. International Journal of Computer Vision, 2021, 129(4): 900 – 920.
- [187] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision [C]//Proceedings of the 38th International Conference on Machine Learning, 2021: 8748 – 8763.
- [188] HUANG X, ZHANG Q, FENG Y, et al. HDR-NeRF: high dynamic range neural radiance fields [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 18377 – 18387.